

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 09:51:25 ; Search time 40 Seconds
(without alignments)
1815.540 Million cell updates/sec

Title: US-09-926-799-1

Perfect score: 2896

Sequence: 1 MMASKADPTNNDGTSGAGQ.....YOLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2896	100.0	545	22 AAB49700	Small round struct
2	1979	68.3	544	22 AAB49703	Small round struct
3	1977.5	68.3	546	22 AAB49702	Small round struct
4	1950	67.3	530	15 AAR50972	Norwalk virus str
5	1938	66.9	530	22 AAB49701	Small round struct
6	1900	65.6	530	15 AAR57091	Small round virus
7	1225	42.3	540	22 AAB49706	Small round struct
8	1178.5	40.7	548	22 AAB49705	Small round struct
9	1174.5	40.6	535	22 AAB49707	Small round struct
10	1172.5	40.5	542	22 AAB49708	Small round struct

11	1157.5	40.0	539	22 AAB49704	Small round struct
12	1157	40.0	548	23 AAU91272	Norwalk virus asso
13	1154.5	39.9	550	22 AAB49709	Small round struct
14	1013	35.0	541	22 AAB49710	Small round struct
15	321	11.1	579	18 AAU08143	RUPV capsid protei
16	276	9.5	547	22 AAM50108	Feline calicivirus
17	276	9.5	671	22 AAM50107	Feline calicivirus
18	272.5	9.4	669	22 AAB67461	Amino acid sequenc
19	272	9.4	623	22 AAB67462	Feline Calicivirus
20	272	9.4	668	22 AAB67462	Amino acid sequenc
21	269	9.3	623	22 AAB47043	Feline Calicivirus
22	269	9.3	668	12 AAR10686	Feline calicivirus
23	269	9.3	668	22 AAE04304	Feline calicivirus
24	262.5	9.1	622	22 AAB47045	Feline Calicivirus
25	154	5.3	2206	13 AAR22210	True type 3 poliov
26	147	5.1	40	23 AAU91273	Norwalk virus asso
27	136	4.7	40	23 AAU91274	Norwalk virus asso
28	132.5	4.6	2209	3 AAP20037	Sequence encoded b
29	125.5	4.3	2179	7 AAP60243	Sequence encoding
30	116.5	4.0	1707	22 ABG22165	Novel human diagn
31	112.5	3.9	2164	9 AAP80131	Peptides translate
32	112.5	3.9	2164	9 AAP81045	Sequence of the vi
33	110	3.8	1037	22 ABB71260	Drosophila melanog
34	110	3.8	1234	22 ABB68510	Drosophila melanog
35	110	3.8	2016	22 ABB63911	Drosophila melanog
36	107.5	3.7	2599	21 AAY75098	Neisseria meningit
37	106	3.7	1302	20 AAY42104	Human MEK1 protei
38	106	3.7	1302	20 AAY26235	Human MEK1 protei
39	106	3.7	1493	20 AAY42107	Murine MEK1 prote
40	106	3.7	1493	20 AAY26234	Murine MEK1 prote
41	106	3.7	1593	23 AAM48935	Murine MEK1-2. M
42	105.5	3.6	239	9 AAP81097	Sequence of rhinov
43	105	3.6	1559	21 AAG50494	Arabisopsis thalia
44	105	3.6	1596	21 AAG50493	Arabisopsis thalia
45	105	3.6	1752	21 AAG50492	Arabisopsis thalia

ALIGNMENTS

RESULT 1
AAB49700
ID AAB49700 standard; protein; 545 AA.
XX
AC AAB49700;
XX
XX
DT 04-APR-2001 (first entry)
DE
DE Small round structured virus protein SEQ ID 1.
XX
XX Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
PN WO200079280-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-JP04095.
XX
PR 22-JUN-1999; 95Jp-0175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
PI
DR WPI; 2001-080848/09.
DR N-PSDB; AAF29141.
XX
PT Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies

```

XX PS Claim 1; Page 40-42; 84pp; Japanese.
XX CC This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks.
XX SQ Sequence 545 AA;
XX Query Match 100.0%; Score 2896; DB 22; Length 545;
XX Best Local Similarity 100.0%; Pred. No. 3.1e-250;
XX Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMASKDAPTNDGTSAGOLVPEANTAEPIISMEPVAGAAATAAGQVNMIDPWIMNNY 60
DB 1 MMASKDAPTNDGTSAGOLVPEANTAEPIISMEPVAGAAATAAGQVNMIDPWIMNNY 60
QY 61 VQAPQGEFTISPNTGDIQLGPHLNPFLSHLAQYNGWGNMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNTGDIQLGPHLNPFLSHLAQYNGWGNMKVKVLLAGNAFTA 120
QY 121 GKIIISCIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLFNHNDNAPTMR 180
DB 121 GKIIISCIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLFNHNDNAPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVEQKTFPSVPLNPLN 240
DB 181 LVCMLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVEQKTFPSVPLNPLN 240
QY 241 TLNSRVPSLIKSMVSRDHQMVQFQNGRVTLDGQLGTTPTTSASOLCKIRGSVFHANG 300
DB 241 TLNSRVPSLIKSMVSRDHQMVQFQNGRVTLDGQLGTTPTTSASOLCKIRGSVFHANG 300
QY 301 GNGYNLTDELGSPYHAFESAPIGFDPDLGECDDHMEASPTQNTGDIKQINVKQESAF 360
DB 301 GNGYNLTDELGSPYHAFESAPIGFDPDLGECDDHMEASPTQNTGDIKQINVKQESAF 360
QY 361 APLHGTQADGLSDVSVNTNMIKLGWSPVSDGHRGDDVPWIPRYGSTLTAAQLAPP 420
DB 361 APLHGTQADGLSDVSVNTNMIKLGWSPVSDGHRGDDVPWIPRYGSTLTAAQLAPP 420
QY 421 IYPGFGGAIVFFMSDFPIAHGTNGLSVPCPTIQEFVTHFVNECAPRGEAALLHYLDPD 480
DB 421 IYPGFGGAIVFFMSDFPIAHGTNGLSVPCPTIQEFVTHFVNECAPRGEAALLHYLDPD 480
QY 481 THRNLFGEFKLYPEGFMTCVPNSSGTGPOTLPINGVFVFSWVSRYQLKPVGTAGPACRL 540
DB 481 THRNLFGEFKLYPEGFMTCVPNSSGTGPOTLPINGVFVFSWVSRYQLKPVGTAGPACRL 540
QY 541 GIRRS 545
DB 541 GIRRS 545
RESULT 2
AAB49703
ID AAB49703 standard; protein: 544 AA.
XX AAB49703;
XX AC AAB49703;
XX DT 04-APR-2001 (first entry)
XX DE Small round structured virus protein SEQ ID 4.
XX KW Small round structured virus; SRSV; food poisoning.
XX OS Small round structured virus.
XX PN WO200007940-A1.

```

```

XX 28-DEC-2000.
XX 22-JUN-2000; 2000WO-JP04095.
XX 22-JUN-1999; 99JP-0175928.
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI: 2001-080848/09.
XX N-PSDB; AAB29144.
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies -
XX Claim 1; Page 47-49; 84pp; Japanese.
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks.
XX SQ Sequence 544 AA;
XX Query Match 68.3%; Score 1979; DB 22; Length 544;
XX Best Local Similarity 66.8%; Pred. No. 3.6e-168;
XX Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;
QY 1 MMASKDAPTNDGTSAGOLVPEANTAEPIISMEPVAGAAATAAGQVNMIDPWIMNNY 60
DB 1 MMASKDAPTNDGTSAGOLVPEANTAEPIISMEPVAGAAATAAGQVNMIDPWIMNNY 60
QY 61 VQAPQGEFTISPNTGDIQLGPHLNPFLSHLAQYNGWGNMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNTGDIQLGPHLNPFLSHLAQYNGWGNMKVKVLLAGNAFTA 120
QY 121 GKIIISCIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLFNHNDNAPTMR 180
DB 121 GKIIISCIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLFNHNDNAPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVEQKTFPSVPLNPLN 240
DB 181 LVCMLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVEQKTFPSVPLNPLN 240
QY 241 TLNSRVPSLIKSMVSRDHQMVQFQNGRVTLDGQLGTTPTTSASOLCKIRGSVFHANG 300
DB 241 TLNSRVPSLIKSMVSRDHQMVQFQNGRVTLDGQLGTTPTTSASOLCKIRGSVFHANG 300
QY 301 GNGYNLTDELGSPYHAFESAPIGFDPDLGECDDHMEASPTQNTGDIKQINVKQES 358
DB 301 GNGYNLTDELGSPYHAFESAPIGFDPDLGECDDHMEASPTQNTGDIKQINVKQES 358
QY 359 AFAPHLGTQADGLSDVSVNTNMIKLGWSPVSDGHRGDDVPWIPRYGSTLTAAQLA 418
DB 359 AFAPHLGTQADGLSDVSVNTNMIKLGWSPVSDGHRGDDVPWIPRYGSTLTAAQLA 418
QY 419 PPIYPGFGGAIVFFMSDFPIAHGTNGLS- ---VPCITPQEFVTHFVNECAPRGEAALL 474
DB 419 PPIYPGFGGAIVFFMSDFPIAHGTNGLS- ---VPCITPQEFVTHFVNECAPRGEAALL 474
QY 475 HYLDPDTHRNLFGEFKLYPEGFMTCVPNSSGTGPOTLPINGVFVFSWVSRYQLKPVGTGA 534
DB 475 HYLDPDTHRNLFGEFKLYPEGFMTCVPNSSGTGPOTLPINGVFVFSWVSRYQLKPVGTGA 534
QY 535 GPA-CRUGIRR 544
DB 535 GPA-CRUGIRR 544

```

```

Db 534 GPARGRLGVRR 544
RESULT 3
ID AAB49702 standard; protein; 546 AA.
AC AAB49702;
XX
XX 04-APR-2001 (first entry)
DE Small round structured virus protein SEQ ID 3.
XX
XX Small round structured virus; SRSV; food poisoning.
XX
XX Small round structured virus.
OS
XX WO2000079280-A1.
PN
XX 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP04095.
XX
XX 22-JUN-1999; 95JP-0175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI; 2001-080848/09.
DR N-PSDB; AAF29143.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies -
XX
XX Claim 1; Page 45-47; 84pp; Japanese.
PS
XX This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks.
XX
XX Sequence 546 AA;
SQ
Query Match 68.3%; Score 1977.5; DB 22; Length 546;
Best Local Similarity 66.6%; Pred. No. 4.9e-168;
Matches 367; Conservative 76; Mismatches 95; Indels 13; Gaps 7;
QY 1 MMWASKDAPTNNDGTSGAGQLVPEANTAEPTSMPEVAGAAATAAGQVMNIDPWIMNY 60
Db 1 MMWASKDAQSDAGAGAGQLVPEYNTADPLPMEPVAGTAVATAGQVMNIDPWVNNF 60
QY 61 VOAPQGETISPNNPTGDTLFDLQGLPHLNPFLSHLAQYNGWGNMKVKVLLAGNAFTA 120
Db 61 VQSPQGETISPNNPTGDTLFDLQGLPHLNPFLSHLSQYNGWGNMKVRILLAGNAFSA 120
QY 121 GKIIISCIPPGFAAQNISIAQTAMFPHVIADVRLIEPVLIEDYRNVLFHNNNDAPTMR 180
Db 121 GKIIIVCCVPPGPTSSSLTAQATLPHVIADVRLTEIEMPLEDYRNVLVHTNDNQPTMR 180
QY 181 LVCMLYTLRASGSSGTDTPFVIAGRVLTCPSPDFSFLFVPPNVEQTKPFSVNLPLN 240
Db 181 LVCMLYTLRTCGGSGNDSFVYAGRVLTAPSSDFSFLFVPPTEQKTRAFVTPNPLQ 240
QY 241 TLNSRVRPLSLIKMVSRRDHGMQVQFNGRVTLTGQLOGTTPTSASOLCKIRGSVFHANG 300
Db 241 TLNSRFRFSLIOGMILSPDASQVQVQFNGRCLIDGLLGTTPATSGQLFRVRGKI - 298

```

```

QY 301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGECDMHMEASPT-TQFNTGDVIKOINVKQE-S 358
Db 299 ARTLNLTEVDGKPFMAFDSAPVGPFDGKCDWHRMRSKTPNNTSSGDPMRSVSVOTNVQ 358
QY 359 AFAPHLGTIQADGLSDSVYNTNNIAKLGHVSPVSDGHRGDVDPWVLPYRGSTLTAEAAQLA 418
Db 359 GFVPHLGSIQFDEVFNHPTG-DYIGTIEWISQFSTPPGTDINLWEIPDYGSSLSQAANLA 417
QY 419 PPIYPGGEAIVFFMSDFPIAHGTNGLS-VPCTIPQEFVTHFVNEQAPTREGAALL 474
Db 418 PVFPPGGEALVYFVSAFP--GPNRSAPNDVPCLLPQEVITHFVSEQAPTMGDAALL 474
QY 475 HYLDDPTHNLGEFKLYPEGFTCVNPSNGTGPQTLPIGVFVSWVSRFYQLKPVGTA 534
Db 475 HYVDPTNRNLGEFKLYPGGYLTCVPNGVGAGPQQLPLNGVLEFVSWVSRFYQLKPVGTA 534
QY 535 GPA-CRLGIRR 544
Db 535 STARSRLGVRR 545
RESULT 4
AAR50972
ID AAR50972 standard; protein; 530 AA.
XX
XX AAR50972;
XX
XX 05-OCT-1994 (first entry)
DT
XX Norwalk virus strain 8FIIa protein (encoded by ORF2).
DE
XX Norwalk virus; pathogen; acute gastroenteritis; food poisoning;
KW seafood contamination; diagnostic assay; calicivirus;
KW small round virus.
XX
XX Norwalk virus (strain 8FIIa).
OS
XX WO9405700-A.
PN
XX 17-MAR-1994.
PD
XX 07-SEP-1993; 93WO-US08447.
PF
XX 07-SEP-1992; 92US-0941365.
PR
XX (BAYU ) BAYLOR COLLEGE MEDICINE.
PA
XX Estes MK, Graham DY, Jiang X, Matson DO;
PI
XX WPI; 1994-101125/12.
DR N-PSDB; AAQ56826.
XX
XX DNA from Norwalk and related viruses - used for preparing prods.
PT for use in diagnostic assays, detection and vaccines for Norwalk
PT and related viruses
XX
XX Claim 14; Page 68-70; 156pp; English.
PS
XX The Norwalk virus was isolated from stool samples from adult
CC volunteers infected with safety tested Norwalk virus strain 8FIIa.
CC The coding sequence is useful for the design of probes for use in
CC diagnostic assays for the Norwalk and related viruses.
XX
XX Sequence 530 AA;
SQ
Query Match 67.3%; Score 1950; DB 15; Length 530;
Best Local Similarity 67.0%; Pred. No. 1.3e-165;
Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;
QY 1 MMWASKDAPTNNDGTSGAGQLVPEANTAEPTSMPEVAGAAATAAGQVMNIDPWIMNY 60
Db 1 MMWASKDATSSVDGAGAGQLVPEVNASDPLAMPVAGSSTAVATAGQVNPIDPWIMNF 60

```

QY 61 VQAPQGETTSPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVKVLLAGNAFTA 120
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks.
XX
SQ Sequence 530 AA;
Query Match 66.9%; Score 1938; DB 22; Length 530;
Best Local Similarity 67.0%; Pred. No. 1.6e-164;
Matches 366; Conservative 63; Mismatches 99; Indels 18; Gaps 7;
QY 1 MMASKDAPNTMDGTSAGOLVPEANTAPISNEPVAGATAAATAGOVNMDPIMNNY 60
Db 1 MMASKDATSSVGCAGAGOLVPEVNASDPLAMPVAGSSTAVATAGQVNPDPWINNF 60
QY 61 VQAPQGETTSPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVKVLLAGNAFTA 120
Db 61 VQAPQGETTSPNNTPGCVLFDLSLGLPHLNPFLSHLSQMYNGWGNMRVRLMAGNAFTA 120
QY 121 GKIIISCIPPGFAAQNTSIAQATMFPHVIADRVLEPIEVPLEDVRNVLPHNND-NAPT 179
Db 121 GKIIISCIPPGFSGHNLTIQAATLFPHVIADRVTLDPVLEPVEDVRNVLPHNDRNQQT 180
QY 180 RLVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQKTPFSPVNLPL 239
Db 181 RLVCMLYTPLRTGGTG--DSFVVGVRVMTCPSPDFNLFVLPPTVEQKTRPFTLNLPL 238
QY 240 NTLNSRVPSLIKSMVSRDHQMVQFQNGRVTLDGQLQGTTPTSASOLCKIRGSVFHAN 299
Db 239 SLSLSNRAPLPISSMGISPDNVQSVQFQNGRCTLDGRVLGVTTPVSLSHVAKIRGT---S 294
QY 300 GNGYNLTGDSGYHAFESAPIGFDPDLGECDMHMEASPTQFNTGDIKQINVKQESA 359
Db 295 NGTVINLTGDTFPHFEGPAPIGFDPDLGGCDWHIN---MTQFGHSSQTQYDVTTPDT 351
QY 360 FAPHLGTIQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPPWIPRYGSTLITEAAQLAP 419
Db 352 FVPHLGSIOANGIG---SGNYIGVLSWSPSPHSPSGSQVDLWKIPNYGSSITEATHLAP 407
QY 420 PIYPGFGAEIVFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNEQAPTRGEAALLHYLDP 479
Db 408 SVYPPGFEVLVFFEMSKMP---GPGAYNLPCLLPOEYIISHLASEQAPTVEAALLHYVDP 464
QY 480 DTHRNLTGFEKLYPEGFMTCVPNSSGTGPQTLPINGVFVFSWSRFFYQLKPVGTAGPA-C 538
Db 465 DTGRNLGEFKAYPDGFLTCVPNGASSGPPQLPINGVFVFSWSRFFYQLKPVGTASSARG 524
QY 539 RLGLRR 544
Db 525 RLGLRR 530
RESULT 5
AAB49701
ID AAB49701 standard; protein; 530 AA.
XX
AC AAB49701;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 2.
XX
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
PN WO200079280-A1.
XX
PD 28-DEC-2000.
XX
PE 22-JUN-2000; 2000WO-JP04095.
XX
PR 22-JUN-1999; 99JP-0175928.
XX
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX
PA (DENK-) DENKA SEIKEN KK.
XX
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI; 2001-080848/09.
DR N-PSDB; AAF29142.
XX
PT Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies
XX
PS Claim 1; Page 42-45; 84pp; Japanese.
XX

CC This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks.
XX
SQ Sequence 530 AA;
Query Match 66.9%; Score 1938; DB 22; Length 530;
Best Local Similarity 67.0%; Pred. No. 1.6e-164;
Matches 366; Conservative 63; Mismatches 99; Indels 18; Gaps 7;
QY 1 MMASKDAPNTMDGTSAGOLVPEANTAPISNEPVAGATAAATAGOVNMDPIMNNY 60
Db 1 MMASKDATSSVGCAGAGOLVPEVNASDPLAMPVAGSSTAVATAGQVNPDPWINNF 60
QY 61 VQAPQGETTSPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVKVLLAGNAFTA 120
Db 61 VQAPQGETTSPNNTPGCVLFDLSLGLPHLNPFLSHLSQMYNGWGNMRVRLMAGNAFTA 120
QY 121 GKIIISCIPPGFAAQNTSIAQATMFPHVIADRVLEPIEVPLEDVRNVLPHNND-NAPT 179
Db 121 GKIIISCIPPGFSGHNLTIQAATLFPHVIADRVTLDPVLEPVEDVRNVLPHNDRNQQT 180
QY 180 RLVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQKTPFSPVNLPL 239
Db 181 RLVCMLYTPLRTGGTG--DSFVVGVRVMTCPSPDFNLFVLPPTVEQKTRPFTLNLPL 238
QY 240 NTLNSRVPSLIKSMVSRDHQMVQFQNGRVTLDGQLQGTTPTSASOLCKIRGSVFHAN 299
Db 239 SLSLSNRAPLPISSMGISPDNVQSVQFQNGRCTLDGRVLGVTTPVSLSHVAKIRGT---S 294
QY 300 GNGYNLTGDSGYHAFESAPIGFDPDLGECDMHMEASPTQFNTGDIKQINVKQESA 359
Db 295 NGTVINLTGDTFPHFEGPAPIGFDPDLGGCDWHIN---MTQFGHSSQTQYDVTTPDT 351
QY 360 FAPHLGTIQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPPWIPRYGSTLITEAAQLAP 419
Db 352 FVPHLGSIOANGIG---SGNYIGVLSWSPSPHSPSGSQVDLWKIPNYGSSITEATHLAP 407
QY 420 PIYPGFGAEIVFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNEQAPTRGEAALLHYLDP 479
Db 408 SVYPPGFEVLVFFEMSKIP---GPGAYSLPCLLPOEYIISHLASEQAPTVEAALLHYVDP 464
QY 480 DTHRNLTGFEKLYPEGFMTCVPNSSGTGPQTLPINGVFVFSWSRFFYQLKPVGTAGPA-C 538
Db 465 DTGRNLGEFKAYPDGFLTCVPNGASSGPPQLPINGVFVFSWSRFFYQLKPVGTASSARG 524
QY 539 RLGLRR 544
Db 525 RLGLRR 530
RESULT 6
AAR57091
ID AAR57091 standard; Protein; 530 AA.
XX
AC AAR57091;
XX
DT 05-OCT-1994 (first entry)
XX
DE Small round virus SRSV/KY/89 capsid protein.
XX
KW pathogen; acute gastroenteritis; food poisoning;
KW seafood contamination; diagnostic assay; human calicivirus;
KW small round virus; SRSV; KY89; Norwalk virus; capsid protein.
XX
OS Small round virus SRSV/KY/89.
XX
PN WO9405700-A.
XX

PD 17-MAR-1994.
XX
XX PF 07-SEP-1993; 93WO-US08447.
XX
XX PR 07-SEP-1992; 92US-0941365.
XX
XX PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
XX PI Estes MK, Graham DY, Jiang X, Matson DO;
XX
XX WPI; 1994-101125/12.
DR N-PSDB; AAQ56832.
XX
XX DNA from Norwalk and related viruses - used for preparing prods.
PT for use in diagnostic assays, detection and vaccines for Norwalk
PT and related viruses
XX
XX Example 7; Fig 13a; 156pp; English.
XX
XX The known sequence for Norwalk virus was used to obtain the
CC sequence of other Norwalk-related viruses such as SRSV/KY/89, an
CC agent from a stool from an outbreak of gastroenteritis in Japan
CC in 1989. The 2516 nucleotide cDNA sequence includes part of the
CC polymerase region and the capsid region of the genome; the deduced
CC amino acid sequences are AA37092 and AA37091, respectively. Expression
CC of fragments and derivs. of Norwalk-related viruses permits
CC development of diagnostic assays to detect antibodies, antigens,
CC viral genetic material or antivirals.
XX
XX Sequence 530 AA;
SQ

Query Match 65.6%; Score 1900; DB 15; Length 530;
Best Local Similarity 65.7%; Pred. No. 4e-161;
Matches 362; Conservative 66; Mismatches 95; Indels 28; Gaps 9;

QY 1 MMWASKDAPTNDGTSAGQOLVPEANTAEPISEMPVAGATAAATAGQVMMDPWTMNNY 60
Db 1 MMWASKDAPTNDGTSAGQOLVPEANTAEPISEMPVAGATAAATAGQVMMDPWTMNNY 60

QY 61 VQAPGEEFTISPNNTPGDILDLQGLPHLNPFLSHLAQMYNGWGNKVKVLLAGNAFTA 120
Db 61 VQAPGEEFTISPNNTPGDILDLQGLPHLNPFLSHLAQMYNGWGNKVKVLLAGNAFTA 120

QY 121 GKIIISCIPPPGAAQNIATAQATMFPHTVADRVLEPIEVPLEDVRNVLHNN-D-NAPTM 179
Db 121 GKIIISCIPPPGAAQNIATAQATMFPHTVADRVLEPIEVPLEDVRNVLHNNDRNQTM 180

QY 180 RLVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSEFLFVPPNVEQTKPFSVNLPL 239
Db 181 RLVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSEFLFVPPNVEQTKPFSVNLPL 238

QY 240 NTLNSRVPPLIKSMWVRDHQMVQFQNGRVTLTGLOGQTTPTSASOLCKIRGVSVFHAN 299
Db 239 SLSLSRAPLPLISGMISPDNVQSQFQNGRCTLGRLVGTTPVSLHSHVAKIRGT----S 294

QY 300 GNGVNLTELDGSPYHAFESPAPIGFPDLGECDDHMEA-----SPTQNTGDVIKQINV 354
Db 295 NCTVNLTELDGTPHPFEGPAPIGFPDLGECDDHNMFTQFGHSSQTOYDV-DTTPDTSV 353

QY 355 KQESAFAPHLGTIOADGLSDVSVNNTMIKLGWSPVSDGHRGDVDPVPIRYGSTLTFA 414
Db 354 -----PHLGSIQANGIG-----SGNYIGVLSWSPSPSPSGQVLDLWKIPNYSITEA 402

QY 415 AQLAPYIPPGFGEATVFMSPDPIAHGTNGLSVPTCTIPQEFVTHFVNBOAPTRGEAALL 474
Db 403 THLAPSVISFGEVLYFPMKIP---GPGGSLPCLLLPGQYISHLASQAQTVGEGPLL 459

QY 475 HYLDPDTHRLNGEFLYPEGFTVCVPNSGCTGPQTLPIGVFVFSWVSRRFYQLKRPVGTA 534
Db 460 HVDVDPDTHRLNGEFLYPEGFTVCVPNSGCTGPQTLPIGVFVFSWVSRRFYQLKRPVGTA 519

QY 535 GRA-CRLGTR 544
+ + + + +

Db 520 STARGRLGLRR 530

RESULT 7
AAB49706
ID AAB49706 standard; protein; 540 AA.
XX
XX AC AAB49706;
XX
XX DT 04-APR-2001 (first entry)
XX
XX DE Small round structured virus protein SEQ ID 7.
XX
XX KW Small round structured virus; SRSV; food poisoning.
XX
XX OS Small round structured virus.
XX
XX PN WO200079280-A1.
XX
XX PD 28-DEC-2000.
XX
XX PF 22-JUN-2000; 2000WO-JP04095.
XX
XX XX 22-JUN-1999; 95JP-0175928.
XX
XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX
XX PA (DENK-) DENKA SEIKEN KK.
XX
XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI; 2001-080848/09.
DR N-PSDB; AAF29147.
XX
XX PT Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies
XX
XX PS Claim 1; Page 54-57; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks.
XX
XX Sequence 540 AA;
SQ

Query Match 42.3%; Score 1225; DB 22; Length 540;
Best Local Similarity 46.7%; Pred. No. 1e-100;
Matches 261; Conservative 89; Mismatches 173; Indels 36; Gaps 14;

QY 1 MMWASKDAPTNDGTSAGQOLVPEANTAEPISEMPVAGATAAATAGQVMMDPWTMNNY 60
Db 1 MMWASKDAPTNDGTSAGQOLVPEANTAEPISEMPVAGATAAATAGQVMMDPWTMNNY 60

QY 61 VQAPGEEFTISPNNTPGDILDLQGLPHLNPFLSHLAQMYNGWGNKVKVLLAGNAFTA 120
Db 57 VQAPGEEFTISPNNTPGDILDLQGLPHLNPFLSHLAQMYNGWGNKVKVLLAGNAFTA 116

QY 121 GKIIISCIPPPGAAQNIATAQATMFPHTVADRVLEPIEVPLEDVRNVLH-NNDNAPTM 179
Db 117 GKIIISCIPPPGAAQNIATAQATMFPHTVADRVLEPIEVPLEDVRNVLH-NNDNAPTM 176

QY 180 RLVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSEFLFVPPNVEQTKPFSVNLPL 239
Db 177 RLVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSEFLFVPPNVEQTKPFSVNLPL 234

QY 240 NTLNSRVPPLIKSMWVRDHQMVQFQNGRVTLTGLOGQTTPTSASOLCKIRGVSVFHAN 299
Db 235 GELNSRFLPSIDEMVTSIPNESIVVQPNQGRVTLTGLOGQTTPTSASOLCKIRGVSVFHAN 294


```
PS Claim 1: Page 57-59; 84pp; Japanese.
XX
CC This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks.
XX
SQ Sequence 535 AA;

Query Match 40.6%; Score 1174.5; DB 22; Length 535;
Best Local Similarity 45.0%; Pred. No. 3.4e-96;
Matches 251; Conservative 87; Mismatches 181; Indels 39; Gaps 13;

Qy 1 MMASKDAPTNNDGTSAGOLVPEANTAEPISEMPEVAGAAATAAGOVNMIDPWIMNNY 60
Db 1 MMASNDAAPSNDGAG---LVPEANN-ETWALEPVGASIAAPLTGQNNIIDPWIRUNF 56
Qy 61 VOAPOGFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVKVLLAGNAFTA 120
Db 57 VOAPNGEFTVSPRNSPGVLLNLELGLPDLNPLYLAHLSRMVNGYAGVGVQVLLAGNAFTA 116
Qy 121 GKIIISCIPGFAAQNISIAQATWPHVIADVRLPEIEVPLEDVRNVLFH-NNDNAPTM 179
Db 117 GKLFVAAVPHHPLFENISPGQITWPHVVIDVRLPEVLLPLPDVRNFFHYNQONEPRM 176
Qy 180 RLVCMLYPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVPEOKTKPFVSNPLP 239
Db 177 RLVMALYPLRSNG--SGDDVFTVSCRVLTRPSDFDNLVPPPLESKTKPFTLPILTI 234
Qy 240 NTLNSRVPSLIKSMVSRDHQMVQFNGRVTLDGQLQGTTPTSASQLCKIRGSVFHA 299
Db 235 GELTNSRFPVIDELYTSPNESLVVQPNQRCALDGLQGTTLQLLPTAICSEFRGINQV 294
Qy 300 GGNGY----NITELDGSYPH-AFESPAPIGFPD-----LGECDWMEASPTQFNWG 346
Db 295 SGENHVNNQVNTINGTFDPTGDVPAPLPTPDFSGKLGFLVSQRD-HDNAC-----RSH 348
Qy 347 DVIKOINVQBSAFAPHLTGTTQADGLSDVSVNTNNIAKLGVSPVSDCHRGDVPDWIPR 406
Db 349 DAVIATN---SAKTPKLGATQIGTWEEDDVHINQPTKF---TPVGLFENEGFNQWTLPN 402
Qy 407 YGSLTEAAQLAPPIYPGFGAIVFFMSDFPIAHGTNGLSVPCCTIPQEFVTHFVNEQAP 466
Db 403 YSGALTNLNGLAPPVAPTFPGEQILFFRSHIPLKGGVADPVIDCLLPQEWIQLHVESAP 462
Qy 467 TRGEAALLHYLDPDTHRNLGFEKLYPEGFMTCVNPSSGTGQTLPINGVVFVSVWSRFY 526
Db 463 SOSDVALIRFTNPDTGRVLFEAKLHRSGITVA--NTGSRPIVVPANGYFRFDTWVNOFY 520
Qy 527 QLKPVGTAGPACRLGIRR 544
Db 521 SLAPMGTCN-----GRRR 533

RESULT 10
AAB49708
ID AAB49708 standard; protein; 542 AA.
XX
XX AAB49708;
AC AAB49708;
XX
XX 04-APR-2001 (first entry)
XX
XX Small round structured virus protein SEQ ID 9.
XX
XX Small round structured virus; SRSV; food poisoning.
XX
XX Small round structured virus.
XX
XX W0200079280-A1.
XX
```

```
PD 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP04095.
XX
XX 22-JUN-1999; 99JP-0175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI: 2001-080848/09.
XX N-PSDB; AAF29149.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies -
XX
XX Claim 1: Page 59-61; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks.
XX
XX Sequence 542 AA;

Query Match 40.5%; Score 1172.5; DB 22; Length 542;
Best Local Similarity 46.4%; Pred. No. 5.3e-96;
Matches 257; Conservative 85; Mismatches 185; Indels 27; Gaps 15;

Qy 1 MMASKD-APTNDWDTSGAGOLVPEANTAEPISEMPEVAGAAATAAGOVNMIDPWIMNN 59
Db 1 MMASNDAAPSN----DGAASLVPE-GINETMPLPEVAGASIAAPVAGOTNIDPWIRTN 55
Qy 60 YVQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVKVLLAGNAFT 119
Db 56 FVQAPNGEFTVSPRNSPGVLLNLELGLPDLNPLYLAHLSRMVNGYAGVGVQVLLAGNAFT 115
Qy 120 AKGIISCIPGFAAQNISIAQATWPHVIADVRLPEIEVPLEDVRNVLFH-NNDNAPT 178
Db 116 AGKILFAAIPPNFLVDMISPAQITMLPLHIVDVRTLEPIMLPDVNRNVYHFNNQOPR 175
Qy 179 MRLVCMLYPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVPEOKTKPFVSNPLP 238
Db 176 MRLVAMLYPLRSNG--SGDDVFTVSCRVLTRPTPDFEFYLVPPSVESKTKPFTLPILT 233
Qy 239 LNTLSNSRVPSLIKSMVSRDHQMVQFNGRVTLDGQLQGTTPTSASQLCKIRGSVFHA 298
Db 234 ISELNTRFPIEQLYTAPNETNVVQCNCRCTLDGELQGTQLLSAVCFLOQRTVAD 293
Qy 299 NGGN-GYNLTEL---DGSYPH-AFESPAPIGFPDLDGECWMEASPTQFNQGDV--IKQ 351
Db 294 NGDNWDQNLQLYTPNGASYDPTDEVPAPLGTQDPSGLYGLVLTQDNVNVSTGEAKNAK 353
Qy 352 INVQESA-EAPHLGCTIQADGLSDVS--VNTNMIAKLGVSPVSDCHRGDVPDWIPRYG 408
Db 354 IYISTSGKFTPKIGSI---GLHSITEHVHPNQOSRTPVGVVAVD-ENTPFQOWVLPHYA 409
Qy 409 STLTEAAQLAPPIYPGFGAIVFFMSDFPIAHGTNGLS--VPCTIPQEFVTHFVNEQAP 466
Db 410 GSLALNTNLAPAVAPTFFGEOLLFFRSRVPVQVQLOGQDAFIDCLLPQEWNVHVFQEAAP 469
Qy 467 TRGEAALLHYLDPDTHRNLGFEKLYPEGFMTCVNPSSGTGQTLPINGVVFVSVWSRFY 526
Db 470 SOADVALLIRYVNPDTGRVLFEAKLHRSGITV--SHTGAYPLVVPNGHFRFDSWVNOFY 527
Qy 527 QLKPVGTAGPACRL 540
Db 528 SLAPMGTCNGRRRI 541
```


PR	22-JUN-1999;	99JP-0175928.	
XX	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.		
PA	(DENK-) DENKA SEIKEN KK.		
XX	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;		
XX	WPI; 2001-080848/09.		
DR	N-PSDB; AAF29151.		
XX	Kit for the detection and typing of small round-structured virus (SRSV)		
PT	strains for investigation of food poisoning outbreaks, contains		
PT	antibodies		
XX	Claim 1; Page 64-66; 84pp; Japanese.		
PS	This invention relates to a kit for the detection and typing of small		
XX	round structured virus (SRSV) strains. The kit contains antibodies		
CC	directed against peptides represented in sequences AAB49700 - AAB49710,		
CC	which are each SRV strain specific. Polynucleotide sequences AAF20141 -		
CC	AAF20151 represent cDNA encoding the strain specific proteins. The kit is		
CC	used for detecting and typing strains of SRV in order to prevent the		
CC	spread of infection and to examine the epidemiology of outbreaks.		
XX	Sequence 541 AA;		
SQ	Query Match 35.0%; Score 1013; DB 22; Length 541;		
	Best Local Similarity 41.4%; Pred. No. 9.9e-82;		
	Matches 243; Conservative 81; Mismatches 155; Indels 108; Gaps 19;		
QY	1 MKWASKDAPNMDGTSGAGOLYPEANTAPISMEPVAGATAAATAGQVNMIDPMIMNY 60		
Db	1 MKWASNDAPSSDGAAG---LPEINN-EVMPLEPVAGASLATPVVGQNIIDPWIRNF 56		
QY	61 VQAPQGEFTISPNTGDLDFDLQGLHNLPLSHLAQMYNGWGNKVKVLLAGNAFTA 120		
Db	57 VQAPAGEFTVSPNPGIILLDELGDLNPLYLALARMYHAGMGVQIIVLAGNAFTA 116		
QY	121 GKIIISCIPGFAAQNLSTAAQTMFHVITADVRLPIEVPLEDVRNVLFH-NNDNAPTM 179		
Db	117 GKIIFAAIPGFPYENLSPSQTITMCPHVIIDVRQLLEPFLLPMPDIWNFFHYNQGNPKL 176		
QY	180 RLVCMLYTLPRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPNVEQKTFSPVNLPL 239		
Db	177 RLVMALYTLPLRA--NNSGDDVFTVSCRVLTKPSDFEFTFLVPPTVESKTKQFALPKI 234		
QY	240 NTLNSRVPSLIKSMYSRDHQWQFONGRVTLDGLOGTTTPTSASQLCKIRGSVFHAN 299		
Db	235 SEMTNSRFPVPDVMYKTARNENQVQNGRVTLDGELGTLTLLAVNICKFGEVIAKN 294		
QY	300 GG-NGY---NLTELDGSPYHAFE-SPAPIGFPDLGECDDWHM-----EASPTTQ--- 342		
Db	295 GDVRSYRMDMEITNTDTPIDPTEDTGPIGSPDFQILFGVASQRKNQNPATRAHEA 354		
QY	343 -FNTG--DVIKQINVKQESAFAPHL-----GTQADGLSDVSVNTNMIK 384		
Db	355 IINTGGDHLCPQISSEIYLTSPNILRCTNPQLPQSLGRGTILI--RSDNGHCHDMVG- 411		
QY	385 LGWVSPV-----SDGHRGDVDPVPIPRYSGTLTEAAQLAPPYPPGF 426		
Db	412 ---TSPTPTTPWQWRKCRSGNCCSGHRYPV-PVVMNR----- 447		
QY	427 GEATVFFMSDFPIAGTNGLSVPCTIPEQFVTFVNEQAPTREGAALLHLYDDPTHNLG 486		
Db	448 -----VTWIVLSHKSGFSSTRKLQ-----LNLRW-----LIREINPDTGRVLF 488		
QY	487 EFKLYPGFMTCPVNSSTGTQTLPIGCVFVSVSWSRFQLKPVGT 533		
Db	489 EARLHKQGFITVA--HTGDNPVIMPPNGYFRFEAWNQFYSLAPVGT 533		
RESULT 15			
AAW08143			

ID	AAW08143 standard; Protein; 579 AA.
XX	AAW08143;
AC	12-MAR-1997 (first entry)
XX	RHDV capsid protein.
DT	
XX	Canarypox virus; vaccinia virus; ALVAC; NVVAC; vector; attenuation;
XX	calicivirus; capsid; antigen; RHDV; vaccine; diagnosis; therapy.
XX	Rabbit haemorrhagic disease virus strain Saone et Loire.
OS	
XX	Key
XX	Location/Qualifiers
FT	Modified-site 45
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
FT	Modified-site 281
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
FT	Modified-site 308
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
FT	Modified-site 369
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
FT	Modified-site 393
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
FT	Modified-site 430
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
FT	Modified-site 474
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
FT	Modified-site 481
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
FT	Modified-site 502
FT	/label= Glycosylation
FT	/note= "putative N-glycosylation site"
XX	WO9639177-A1.
XX	12-DEC-1996.
XX	03-JUN-1996; 96WO-IB00721.
XX	06-JUN-1995; 95US-0471025.
XX	(VIRO-) VIROGENETICS CORP.
XX	Fischer L, Legros F, Paoletti E;
XX	WPI: 1997-042857/04.
XX	N-PSDB; AAT42749.
PT	Modified recombinant virus contg. exogenous calicivirus DNA - useful
PT	in vaccines and for in vitro proddn. of calicivirus antigens, for
PT	generation of therapeutic or diagnostic antibodies
XX	Example 12; Fig 12a; 150pp; English.
XX	The rabbit haemorrhagic disease virus (RHDV) capsid protein
CC	(AAW08143) amino acid sequence was deduced from the putative capsid
CC	gene (AAT42749) obtd. from RHDV strain Saone et Loire. The capsid
CC	protein can be expressed from novel recombinant, attenuated
CC	vaccinia virus and canarypox virus vectors, esp. based on NVVAC and
CC	ALVAC, useful as vaccines of improved safety. The protein may
CC	also be expressed in vitro and used to generate therapeutic or
CC	diagnostic antibodies.
XX	Sequence 579 AA;
SQ	

```

Query Match      11.1%; Score 321; DB 18; Length 579;
Best Local Similarity 28.3%; Pred. No. 9.2e-20;
Matches 132; Conservative 61; Mismatches 171; Indels 102; Gaps 21;

QY 12 MDGTSAGQLVPPEANTAEPI$-----MEPVAGAAATRA-----ATAG----- 47
Db 1 MEGKARAAPQGEAAGTATTASVPGTTTGDMDPGVATTSVITAENSASATATAGIGPPQ 60

QY 48 QVNMIDPWIHNNYVQAPQGEFTISPNNTPGDILFDLQPLHNPFLSHLAQMYNGWGNM 107
Db 61 QVQOQETWRTNFY---NDVFTWSVADAGSILYTVQHSPPNNPFTAVLSQMYAGWAGM 117

QY 108 KVKVLLAGNAFTAGKIIISCIPGFAAQNISIA---QATMFPHVIADRVRLPIEVPLED 164
Db 118 QRFRIVAGSGVFGRLVAAVIPG-----IEIGPGLEVQFPHVIDARSLEPVTITMPD 172

QY 165 VRNVLFH-NNDNAPTWRVLCMLYTPL--RASGSSSGTDPFVIAGRVLTCPSPDFSLFLV 221
Db 173 LRPNMYHTGDPGLVPTLVLSVYNNLINPFGSTS-----AIQVTVETTPSEDFEVMIR 227

QY 222 PPNVEQKTKPFSVPNLPLNTLSNRVPSLIKSMVSRDHQMVQFO-----NGRVT 272
Db 228 APS--SKTVDSISFAGLLT-----PVLTCVGNDRNNGQIVGLQPVPGGFCSTCRHWN 279

QY 273 LDGLOQCTTPTSASQLCKIRGSVFHANGNGYNLTTEL---DGSPYHAFESP-----APIG 324
Db 280 LNSTGYWSSPRFADIDHRRGSASY-SGNATNVLQFWYANAGS---AIDNPISQVAPDG 335

QY 325 FPDLGCDHMEASPTQFNQNTGDIQINVKQESAFAPHLGTIQADGLSDVSVNTNNIAK 384
Db 336 FPDMSFVPFNGPGIPAAGWVGFAIWN$-----SGAPNVTTVQA-----YE 377

QY 385 LGWVSPVSDGHRGDVDPWIPRYGSTLTAAQ-LAPPIYPPGGEA 429
Db 378 LGPAT-----GAPGNLQP-----TTTSCAQTVAKSIYAVVTGTA 412

```

Search completed: January 16, 2003, 09:54:50
Job time : 43 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 09:54:10 ; Search time 16 Seconds
(without alignments)
1002.218 Million cell updates/sec

Title: US-09-926-799-1
Perfect score: 2896
Sequence: 1 MMASKDAPTNDGTSGAGQ.....YOLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2.6/ptodata/1/1aa/5A-COMB.pep.*
2: /cgn2.6/ptodata/1/1aa/5B-COMB.pep.*
3: /cgn2.6/ptodata/1/1aa/6A-COMB.pep.*
4: /cgn2.6/ptodata/1/1aa/6B-COMB.pep.*
5: /cgn2.6/ptodata/1/1aa/PCTUS-COMB.pep.*
6: /cgn2.6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	285	9.8	626	4	US-09-590-020-7
2	272	9.4	623	4	US-09-590-020-4
3	269	9.3	623	4	US-09-590-020-2
4	262.5	9.1	622	4	US-09-590-020-6
5	149	5.1	2206	1	US-07-852-260-2
6	149	5.1	2206	2	US-08-461-503-2
7	149	5.1	2206	4	US-08-465-250-2
8	106	3.7	1302	4	US-09-423-890-2
9	106	3.7	1493	4	US-09-423-890-8
10	106	3.7	1593	4	US-08-628-829-4
11	106	3.7	2318	4	US-09-091-219-24
12	103.5	3.6	1170	4	US-09-749-588-2
13	102.5	3.5	588	1	US-08-460-860-4
14	102	3.5	1048	4	US-09-171-699-10
15	100	3.5	544	4	US-09-615-192A-349
16	98	3.4	1209	4	US-09-749-588-4
17	97	3.3	956	4	US-09-134-07-63
18	97	3.3	3443	2	US-08-416-603-2
19	95.5	3.3	202	2	US-08-416-603-10
20	95	3.3	2232	4	US-09-091-219-25
21	95	3.3	2247	4	US-09-091-219-2
22	94.5	3.3	2004	1	US-08-375-709-15
23	94.5	3.3	2004	1	US-08-752-929-15
24	94.5	3.3	2004	1	US-09-090-793-9
25	94	3.2	1036	4	US-09-206-943-73
26	94	3.2	1477	1	US-08-038-682-4
27	94	3.2	1477	1	US-08-302-832-4

28	94	3.2	1477	2	US-08-530-198-4	Sequence 4, Appl
29	94	3.2	1477	2	US-08-469-880-4	Sequence 4, Appl
30	94	3.2	1477	2	US-08-728-470-4	Sequence 4, Appl
31	94	3.2	1477	2	US-08-617-697-4	Sequence 4, Appl
32	94	3.2	1477	4	US-08-719-641-4	Sequence 4, Appl
33	94	3.2	1477	4	US-09-206-942-71	Sequence 71, Appl
34	93.5	3.2	907	3	US-08-783-774-2	Sequence 2, Appl
35	93.5	3.2	907	4	US-09-328-599A-1	Sequence 1, Appl
36	93.5	3.2	907	5	PCT-US95-04611A-19	Sequence 19, Appl
37	93	3.2	20	4	US-08-973-961-5	Sequence 5, Appl
38	93	3.2	1381	4	US-09-540-245A-16	Sequence 16, Appl
39	92.5	3.2	669	4	US-09-071-035-264	Sequence 264, App
40	92.5	3.2	1638	4	US-09-071-035-258	Sequence 258, App
41	92.5	3.2	1638	4	US-09-071-035-262	Sequence 262, App
42	92.5	3.2	1638	4	US-09-071-035-266	Sequence 266, App
43	92.5	3.2	2308	1	US-08-015-973-1	Sequence 1, Appl
44	92.5	3.2	2308	2	US-08-448-164-1	Sequence 1, Appl
45	92.5	3.2	2308	4	US-08-081-929-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-09-590-020-7
; Sequence 7, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnais, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-7

Query Match 9.8%; Score 285; DB 4; Length 626;

Best Local Similarity 23.8%; Pred. No. 4.7e-20;

Matches 144; Conservative 87; Mismatches 194; Indels 180; Gaps 32;

QY	13	DGTSGAGQLVPEANT-AEPISEMPVAGAAATAAGOVNNDPWNNVYQAQGEFTI-70	
DB	82	DGSIITA-----PEOGTAVGGVIAEPSAQMSTAADMATCKSVDSW-----EAFESFH128	
QY	71	-----SPNNTPGDILDLQGLPHLNPFLSHLAQMYGNWGNMKVKVLLAGNAFTAGKIII125	
DB	129	TSVNMSTSETQGRILFKQSLGPLNLPYLSHLAKLYVAMSGSIEVRSISGSGVFGKLA188	
QY	126	SCIPPGF-RAQNTISIAQAQMFPHVIADRVRLVLEPIEPLVDVRLVFNHNDNATMRLVCM184	
DB	189	IYVPPGVDVPQVQSTSMIQ---YPHVLEFDAQRQVEPIFCLPDLRSTLYHLKMSDOTTSLVIM245	
QY	185	LY-----TPLRASGSSGCTDPFVIAGRVLTCPSPDFELFLVPPNVQKTKPFVSPNPLN240	
DB	246	VYNDLNPVANDTNSGSC---IVT--VETKPGDPDFEHLKPPG-----S285	
QY	241	TLSNRPVS-LI---KSMVSRDH-QQMVQFQNGRVTLQGLQGTTPTSASQLCKINGSV295	
DB	286	MLTHGVSVPDLIPKSSSLWGNRHSWDITDF-----IIRPFV332	
QY	296	PHANGNGYN-----LTELDCSPYHA-----FESPA-PIGPPDL---G329	
DB	323	FOANRHFEDFNQETAGWSTFRFRPISVTITEQNGAKLIGVATDYIVPGIPDGWPDPTTIPG382	

QY 306 LTELDSGYHAFESPA-PIGFPDLGEC-----WHMEASPTQF-----NT 345
 Db 400 ITTAAG--YDAAEITVNTNFKSMYICGSLQRAWGDKKISNTAFITTAVRKGNSEPSWT 457
 QY 346 GDVTKQINVKQESAFAPHLGTIQADGLSDYSVNTNMTAKLGWVSPVSDGHRGDVDP----401
 Db 458 IDMTKLV-VYQDA---HVG-----EEVQTSIDITLALLGYTGIGEAIGSDRDKVRI 505
 QY 402 WIPRYGSTLTEAAQLAPPYPPGFGGAIVFFMSDFPIAHGTNGLSVPCPTIPOEFVTHFV 461
 Db 506 SVLPETGAR-----GGNHPIFYKNSIKLGYVIRSIDV-----FN 539
 QY 462 NEQAPTRGEAALLHYLDP-----DTHRNLFGEKLYPEGFMTCPVNSSSGTGOTLP 511
 Db 540 QILHTSRQLSNLHYLLPPDSFAVYRIIDNSGWFIDIGIDTDFG-----SEFVGSNLP 592

RESULT 3

US-09-590-020-2
 ; Sequence 2, Application US/09590020
 ; Patent No. 6355246
 ; GENERAL INFORMATION:
 ; APPLICANT: Kruger, John M
 ; APPLICANT: Maes, Roger K
 ; APPLICANT: Vilnis, Alvars
 ; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
 ; TITLE OF INVENTION: CALCIVIRUS AND VACCINES THEREOF
 ; FILE REFERENCE: MSU 4.1-446
 ; CURRENT APPLICATION NUMBER: US/09/590,020
 ; CURRENT FILING DATE: 2000-06-08
 ; PRIOR APPLICATION NUMBER: 60/138,484
 ; PRIOR FILING DATE: 1999-06-10
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 623
 ; TYPE: PRT
 ; ORGANISM: Feline calicivirus
 US-09-590-020-2

Query Match 9.3%; Score 269; DB 4; Length 623;

Best Local Similarity 29.8%; Pred. No. 2e-18;
 Matches 76; Conservative 41; Mismatches 94; Indels 44; Gaps 9;

QY 7 DAPT-NMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAATAGOVNMIDPWIMNNYVQAPQ 65
 Db 72 DULPRLGDDGSIITPEQGTWVGVIAPSAQMSAAADMATGKSVSEW-----E 122
 QY 66 GEFTI-----SPNTPGDIILDLQGLPHLNPFLSHLAQMYNGWGNMKVLLAGNAFT 119
 Db 123 AFSEHTSVNWTSETQGGKILFKQSLGPNLNPYLSHLAKLYVWVSGSVEVRSISGSGVF 182
 QY 120 AGKIISCIPPGF-AAQNISIAQATMFPHVIADVRVLEPIEVLDPEDVRNLFHNNDAFT 178
 Db 183 GGLAAIVVPPGIEPVQSTMLQ---YPHVLFDAQVEPVIFAIPLDRSLNLYHLMSTDT 239
 QY 179 MRLVCMY-----TPLRASGSSGTPDFVIAGRVLCPSDFSLFLVPPNVQKTKPSV 234
 Db 240 TSLVMVYNDLINPYANDTNSG---IVT--VETKPGDPDFKHLKPPG-----284
 QY 235 PNLPLNTLSNRVPS 249
 Db 285 -----SMLTHGVSVP 294

RESULT 4

US-09-590-020-6
 ; Sequence 6, Application US/09590020
 ; Patent No. 6355246
 ; GENERAL INFORMATION:
 ; APPLICANT: Kruger, John M
 ; APPLICANT: Maes, Roger K
 ; APPLICANT: Vilnis, Alvars

QY 330 EC-----DWHM-----EASPTTQFNTGDVVK-QINVK-----QESAPA 361
 Db 383 ELIPAGDYAITNGTGDITATCYDTADIIKNTNFRGMYICGSLQRAWGDKKISNTAF- 441
 QY 362 PHLGTIQADGLSDYSVN-TNMI-----AKLGWSPVSD-----GHRGDVDPWVI 404
 Db 442 --ITTATLGDGNNKINPCNTIDQSKIVVFODAHVGKKAQTSDDTLALLGYTGIGEQAI- 498
 QY 405 PRYGSTLTEAAQLA--PPIYPPGFGGAIVFFMSDRPIAHGTNGLSVPCPTIPOEFVTHFVN 462
 Db 499 ---GSDRDRVRLSTLPETGARG-GNHPIFYKNSIKLGYVIRSIDV-----FNS 543
 QY 463 EQAPTRGEAALLHYLDP-----DTHRNLFGEKLYPEGFMTCPVNSSSGTGOTLPDI 512
 Db 544 QILHTSRQLSNLHYLLPPDSFAVYRIIDNSGWFIDIGIDSGF--SFVGVSGFGKLEFPL 601

QY 513 NGVEFV 517

Db 602 SASYM 606

RESULT 2

US-09-590-020-4
 ; Sequence 4, Application US/09590020
 ; Patent No. 6355246
 ; GENERAL INFORMATION:
 ; APPLICANT: Kruger, John M
 ; APPLICANT: Maes, Roger K
 ; APPLICANT: Vilnis, Alvars
 ; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
 ; TITLE OF INVENTION: CALCIVIRUS AND VACCINES THEREOF
 ; FILE REFERENCE: MSU 4.1-446
 ; CURRENT APPLICATION NUMBER: US/09/590,020
 ; CURRENT FILING DATE: 2000-06-08
 ; PRIOR APPLICATION NUMBER: 60/138,484
 ; PRIOR FILING DATE: 1999-06-10
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 623
 ; TYPE: PRT
 ; ORGANISM: Feline calicivirus
 US-09-590-020-4

Query Match 9.4%; Score 272; DB 4; Length 623;

Best Local Similarity 22.2%; Pred. No. 9.7e-19;
 Matches 133; Conservative 84; Mismatches 209; Indels 174; Gaps 27;

QY 7 DAPT-NMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAATAGOVNMIDPWIMNNYVQAPQ 65
 Db 72 DULPRLGDDGSIITPEQGTWVGVIAPSAQMSAAADMATGKSVSEW-----E 122
 QY 66 GEFTI-----SPNTPGDIILDLQGLPHLNPFLSHLAQMYNGWGNMKVLLAGNAFT 119
 Db 123 AFSEHTSVNWTSETQGGKILFKQSLGPNLNPYLSHLAKLYVWVSGSVEVRSISGSGVF 182
 QY 120 AGKIISCIPPGF-AAQNISIAQATMFPHVIADVRVLEPIEVLDPEDVRNLFHNNDAFT 178
 Db 183 GGLAAIVVPPGIEPVQSTMLQ---YPHVLFDAQVEPVIFAIPLDRSLNLYHLMSTDT 239
 QY 179 MRLVCMY-----TPLRASGSSGTPDFVIAGRVLCPSDFSLFLVPPNVQKTKPSV 234
 Db 240 TSLVMVYNDLINPYANDTNSG---IVT--VETKPGDPDFKHLKPPG-----284
 QY 235 PNLPLNTLSNRVPS-LI---KSMVMVRDGHQMVQ-----FQ-NGRVTLDGQLQ-- 279
 Db 285 -----SMLTHGVSVPDLIPKSSSLIGNRHWSDITDFTIRPFVFOQNRHFDNFOETAGWS 339
 QY 280 -----ITPTSASOLCKI-----RGSVFHANGNGYN 305
 Db 340 TPRFRPITTVSESNNKLGICGIVATDIYVGPIDGWPDTTIPQLTPAGIYSITASNTGD 399

```

RESULT 5
US-07-852-260-2
; Sequence 2, Application US/07852260
; Patent No. 5525715
; GENERAL INFORMATION:
; APPLICANT: Racanello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham

```

```

15  TSGAQGLVPEANTAEPTSMEDVAGAAATAATAGOV--NM-----IDPTMNNVYVQAPOG-- 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
347 TPGSNQYLTSNDHQSPCAI--PEFDVTPPIDPGEVKNNMELAEIDTMIPLNLESTKRNTM 405
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
67  --EFTTSPNTPGDIILFDLQGLPHLPFLSH-----LAOMYNGWGNKVKVLLAGNAF 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
406 DMYRVTUSDADLSQPLICUSLSFAPDRLSHTMLGGEVLNYYTHWAGSLKFTFLFCGSM 465
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
119 TAGRIISCIPPGFAAQNISIAQTMPHFVIAADVRLPEPIEVLDDVRNVLFH--NNDNA 176
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
466 ATGAILVAYAPPG--AQPTSRKEAMLGTHVLINDLGLOSSCTMVVPWISNVYRQTQDSF 524
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
177 PTMLVCMVY-----TPLRAGSSSGTDPFVIAGRVLTCPSPDPSFLEL-----V 221
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
525 TEGGYISMFQYTRIWWPLSTPKMS-----MLGFVSAC--NDFSRLLRDTHHSQSAL 576
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
222 PPNVEQRTK-----PFSVPNLP-----LNTLSNRVPSLIKSMM 255
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
577 PQGIEDTSEVAQALTLSLPKQDLSLPTDKASGPAHSKEVPAULTAVETGATNPLAPSDT 636
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
256 VSRDHGMQVQ-----FQNGRVTLDGQLGQTTPTSASOLCKTRGSVFII----- 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
637 VQTRH--VVQRRSRSESTIESFFARGACVAILIEDNEQPTTQAOKLFAMWRITYKDTVOL 694
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
298 -----ANGNGYNLTDLGSPYHATESPAPIGFPDLGECDW 333
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
695 RRKLEFFTYSRDEMTFVVVTANFTNANHGALNQVQIYIPPCAPTPKSWDDY---TW 751
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
334 HMEASPTQFNTGDVIRKQINVKQESAFAPHLGTTQA-----DGLSDSVSVNTMIAKLG 386
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
752 QTSSNPISIFTYGAAPARISV-----PYVGLANAYSHEFYDGFAPKVPLEKTDANDQIG 802
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
US-08-461-503-2
; Sequence 2, Application US/08461503
; Patent No. 5834302
; GENERAL INFORMATION:
; APPLICANT: Racanietello, Vincent

```

RESULT 6
US-08-461-503-2
; Sequence 2, Application US/08461503
; Patent No. 5834302
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent

APPLICANT: Tatem, Joanne M.
APPLICANT: Weeks-Levy, Carolyn L.
TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES
TITLE OF INVENTION: FROM CDNA
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,503
FILING DATE: 5-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 36607-D-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX: 423523 COOP UI
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2206 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-503-2

Query Match 5.1%; Score 149; DB 2; Length 2206;
Best Local Similarity 20.3%; Pred. No. 2.5e-05;
Matches 97; Conservative 57; Mismatches 196; Indels 128; Gaps 19;

QY 15 TSGAGOLVPEANTAEPISMEPVAGATAAATAGQV-NM-----IDPWIMNNYVQAPOG-- 66
DB 347 TPGSNQYLTSNDHOSPCAI-PEFDVTPPIDPGEVKNMELAEIDTMIPLNLESTKRNTM 405
QY 67 ----EFTISPNTPGDILFDLQGLPHLNPFLSH-----LAQYNGWGVGNMKVYLLAGNAF 118
DB 406 DMRYVTLSDSADLSQPILCLSLSPAFDPRLSHTMLGEVLYNYTHWAGSLKFTFLFCGSM 465
QY 119 TAGKIIISCIIPGFAQONISIAQATMFPVIAADVRLPIEVPLEDVRNVLPH--NNDNA 176
DB 466 ATGKILVAYAPPG-AQPPTSREKAMLGTHVWDLGLQSSCTWVVPWISNVITYRQTQDSF 524
QY 177 PTMRLVCMLY-----TPLRASGSSGTDPFVIAGRVLTCPSPDPSFLFL-----V 221
DB 525 TEGGYISMFYQTRIVVPLSTPKSMS-----MLGFVSAC--NDFSVRLLRDTTHISQSAL 576
QY 222 PPNVEQTK-----PFSVNLPL-----LNTLSNRVPSLIKSM 255
DB 577 PQGIEDLTSEVAGALTLSLPKQDLSLPTDKASGPAHSKEVPALTAETGATNPLAPSDT 636
QY 256 VSRDHGQMWQ-----FQNGRVTLGQLOQTTPTSASQCKIRGSVFH----- 297
DB 637 VQTRH--VVQRRSRSESTIESFFARGACVAIIIEVDNEQPTTRAQKLFAMWRITYKDTVOL 694
QY 298 -----ANGNGYNLTDLGSPYHAFESPAPIGFPDLGECDW 333
DB 695 RRKLEFFYSRFDMEFTVVVTANFNANNGHALNQVQIMYIPPGAPTPKSMDDY---TW 751
QY 334 HMEASPTQNTGDVIKQINVKQESAFAPHLGTIQA-----DGLSDVSNTNNIAKLG 386
DB 752 QTSNSRSIFTYCAAPARISV-----PYVGLANAYSHFYDGFAPVPLKTDANDQIG 802

RESULT 7
US-08-465-250-2
; Sequence 2, Application US/08465250
; Patent No. 6136570
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release 1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,250
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-E-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-465-250-2

Query Match 5.1%; Score 149; DB 4; Length 2206;
Best Local Similarity 20.3%; Pred. No. 2.5e-05;
Matches 97; Conservative 57; Mismatches 196; Indels 128; Gaps 19;

QY 15 TSGAGOLVPEANTAEPISMEPVAGATAAATAGQV-NM-----IDPWIMNNYVQAPOG-- 66
DB 347 TPGSNQYLTSNDHOSPCAI-PEFDVTPPIDPGEVKNMELAEIDTMIPLNLESTKRNTM 405
QY 67 ----EFTISPNTPGDILFDLQGLPHLNPFLSH-----LAQYNGWGVGNMKVYLLAGNAF 118
DB 406 DMRYVTLSDSADLSQPILCLSLSPAFDPRLSHTMLGEVLYNYTHWAGSLKFTFLFCGSM 465
QY 119 TAGKIIISCIIPGFAQONISIAQATMFPVIAADVRLPIEVPLEDVRNVLPH--NNDNA 176
DB 466 ATGKILVAYAPPG-AQPPTSREKAMLGTHVWDLGLQSSCTWVVPWISNVITYRQTQDSF 524
QY 177 PTMRLVCMLY-----TPLRASGSSGTDPFVIAGRVLTCPSPDPSFLFL-----V 221
DB 525 TEGGYISMFYQTRIVVPLSTPKSMS-----MLGFVSAC--NDFSVRLLRDTTHISQSAL 576
QY 222 PPNVEQTK-----PFSVNLPL-----LNTLSNRVPSLIKSM 255
DB 577 PQGIEDLTSEVAGALTLSLPKQDLSLPTDKASGPAHSKEVPALTAETGATNPLAPSDT 636
QY 256 VSRDHGQMWQ-----FQNGRVTLGQLOQTTPTSASQCKIRGSVFH----- 297
DB 637 VQTRH--VVQRRSRSESTIESFFARGACVAIIIEVDNEQPTTRAQKLFAMWRITYKDTVOL 694
QY 298 -----ANGNGYNLTDLGSPYHAFESPAPIGFPDLGECDW 333


```
Db 695 RRKLEFTYSRDEHTEFTVVTANNGHALNQVQIMYIPPGAPTPKSWDDY---TW 751
QY 334 HMEASPTQFNTGDKVIQINVKQESAFAPHLGTIQ-----DGLSDVSVNTNMIKLG 386
      :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 752 QTSSNPISIFTYGAAPARISV-----PVGLANAYSHEYDGFAPKVPKLTANDQIG 802

RESULT 8
US-09-423-890-2
; Sequence 2, Application US/09423890
; Patent No. 6312934
; GENERAL INFORMATION:
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: HUMAN MEK PROTEIN AND NUCLEIC ACID MOLECULES
; FILE REFERENCE: CPI-085CPCP
; CURRENT APPLICATION NUMBER: US/09/423,890
; CURRENT FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: USSN 60/078,153
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: USSN 60/099,165
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-890-2

Query Match 3.7%; Score 106; DB 4; Length 1302;
Best Local Similarity 18.7%; Pred. No. 0.24;
Matches 83; Conservative 72; Mismatches 165; Indels 124; Gaps 20;

QY 1 MMASKDAPTNMDGTSCAGOLVPEANTAEPISEMPVAG-AATAAATAGQVNMIDPHIMNN 59
      :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 400 LLLANGESTGSGGSG-GSL--SAGAAAGSSQPSISGDVVEACCSVLSTVCADP-VYKV 455

QY 60 YVQA---POGEFTISPNTPGD-ILFDLQGLPHLNPFL-----SHLAQMY 100
      || :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 456 YVAALKTLRAMLVYTPCHSLAEIKLQRLRPVVDILVKCADANSRTSOLSISTVLEIC 515

QY 101 NGWGNMKV--KVLGAGNAFTAG-KTIISCIPGFAAQN-----ISIAQA 142
      || :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 516 NQAGKLVAGREILKAGSIGVGVDVYVLSCLCNQAESNNWQELLGRCLLDRLLEFPA 575

QY 143 TMFPHVIA-DVRVLEPIEVPLEDVRNVL---FHNNDNAPTM--RLVCMLYTPLRASCSS 196
      :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 576 EYPHIVSTDVSOAEPVEIRYKLLSLTLFALQSIDNSHSMWGLKSRRIY----- 625

QY 197 GTPDFVIAGRLVTCPSPDFSFLFVPPNVBEQKTPFSVPNPLPLNTLSNRVPSLIKSMV 256
      :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 626 -----LSSARMVTA VAVFSKLVTM-----LNASGSTHFTMRRLMA 663

QY 257 SRDHGMQVQ-FQNG-RVTLDG----QLQGTTPTSASOLCKIRGSVFHANGNGYNLTELD- 310
      || :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 664 IADEVEIAEVIQLGVEDTVGDHDSLQALAPASCLSENSLSLEHTVHREKTKGLSATRLSA 723

QY 311 -----CSPYHAFESPAPIGFPDPLGECDDHWEAS 338
      || :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 724 SSEDISDRLAGSVGLPSSTTTQPKPAVOTKGRPHSQCLNSSLPLSHAQL-----MPPA 777
      |: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:

QY 339 PTTQFNTGDKVIQINVKQESAFAP 362
      |: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 778 PSAPCSSAPSPVPSIKHRPQAFVP 801

RESULT 9
US-09-423-890-8
; Sequence 8, Application US/09423890
; Patent No. 6312934
; GENERAL INFORMATION:
; APPLICANT: Johnson, Gary L.
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To Exter
; FILE REFERENCE: CPI-004DVC3
; CURRENT APPLICATION NUMBER: US/08/628,829A
; CURRENT FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: 08/440,421
; EARLIER FILING DATE: 1995-05-15
; EARLIER APPLICATION NUMBER: 08/323,460
; EARLIER FILING DATE: 1994-10-14
; EARLIER APPLICATION NUMBER: 08/049,254
; EARLIER FILING DATE: 1993-05-15
; EARLIER APPLICATION NUMBER: 08/410,602
; EARLIER FILING DATE: 1995-04-24
```

```
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: HUMAN MEK PROTEIN AND NUCLEIC ACID MOLECULES
; FILE REFERENCE: CPI-085CPCP
; CURRENT APPLICATION NUMBER: US/09/423,890
; CURRENT FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: USSN 60/078,153
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: USSN 60/099,165
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1493
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-423-890-8

Query Match 3.7%; Score 106; DB 4; Length 1493;
Best Local Similarity 18.7%; Pred. No. 0.31;
Matches 83; Conservative 72; Mismatches 165; Indels 124; Gaps 20;

QY 1 MMASKDAPTNMDGTSCAGOLVPEANTAEPISEMPVAG-AATAAATAGQVNMIDPHIMNN 59
      :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 591 LLLANGESTGSGGSG-GSL--SAGAAAGSSQPSISGDVVEACCSVLSTVCADP-VYKV 646

QY 60 YVQA---POGEFTISPNTPGD-ILFDLQGLPHLNPFL-----SHLAQMY 100
      || :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 647 YVAALKTLRAMLVYTPCHSLAEIKLQRLRPVVDILVKCADANSRTSOLSISTVLEIC 706

QY 101 NGWGNMKV--KVLGAGNAFTAG-KTIISCIPGFAAQN-----ISIAQA 142
      || :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 707 KQAGELAVGRELKAGSIGVGVDVYVLSCLCNQAESNNWQELLGRCLLDRLLEFPA 766

QY 143 TMFPHVIA-DVRVLEPIEVPLEDVRNVL---FHNNDNAPTM--RLVCMLYTPLRASCSS 196
      :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 767 EYPHIVSTDVSOAEPVEIRYKLLSLTLFALQSIDNSHSMWGLKSRRIY----- 816

QY 197 GTPDFVIAGRLVTCPSPDFSFLFVPPNVBEQKTPFSVPNPLPLNTLSNRVPSLIKSMV 256
      :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 817 -----LSSARMVTA VAVFSKLVTM-----LNASGSTHFTMRRLMA 854

QY 257 SRDHGMQVQ-FQNG-RVTLDG----QLQGTTPTSASOLCKIRGSVFHANGNGYNLTELD- 310
      || :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 855 IADEVEIAEVIQLGVEDTVGDHDSLQAVAPTSCLENSLSLEHTVHREKTKGLSATRLSA 914

QY 311 -----CSPYHAFESPAPIGFPDPLGECDDHWEAS 338
      || :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 915 SSEDISDRLAGSVGLPSSTTTQPKPAVOTKGRPHSQCLNSSLPLSHAQL-----MPPA 968
      |: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:

QY 339 PTTQFNTGDKVIQINVKQESAFAP 362
      |: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
Db 969 PSAPCSSAPSPVPSIKHRPQAFVP 992

RESULT 10
US-08-628-829-4
; Sequence 4, Application US/08628829A
; Patent No. 6333170
; GENERAL INFORMATION:
; APPLICANT: Johnson, Gary L.
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To Exter
; FILE REFERENCE: CPI-004DVC3
; CURRENT APPLICATION NUMBER: US/08/628,829A
; CURRENT FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: 08/440,421
; EARLIER FILING DATE: 1995-05-15
; EARLIER APPLICATION NUMBER: 08/323,460
; EARLIER FILING DATE: 1994-10-14
; EARLIER APPLICATION NUMBER: 08/049,254
; EARLIER FILING DATE: 1993-05-15
; EARLIER APPLICATION NUMBER: 08/410,602
; EARLIER FILING DATE: 1995-04-24
```

```

; EARLIER APPLICATION NUMBER: 08/472,934
; EARLIER FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; SEQ ID NO 4
; LENGTH: 1593
; TYPE: PRT
; ORGANISM: Mus musculus
; US-08-628-829-4

Query Match          3.7%; Score 106; DB 4; Length 1593;
Best Local Similarity 18.7%; Pred. No. 0.34;
Matches 83; Conservative 72; Mismatches 165; Indels 124; Gaps 20:

QY 1 MMWASKDAPTNMDGTSGAGOLVPEANTAEPISEPVAG-AATAAATAGOVNMIDPWIMNN 59
Db 691 LLLANGESTNGSGSG-GSL--SAGASGSSQPSISGDVVEACCSVLVCA DP-VYKV 746

QY 60 YVQA---PQGEFTISPNTPGD-ILFDLQLGPHLNPFL-----SHLAQMY 100
Db 747 YVAALKTLRAMLVYTPCHSLAERIKLQRLRPVVDTILVKCADANSRTSOLSISTVLELC 806

QY 101 NGWGNMKV--KVLLAGNATAG-KIISCIPGFRAQN-----ISIAQA 142
Db 807 KGQAGELAVGREILKAGSIGVGGVDYVLSCLIGNQAESNNWQELLGRCLIDRLLEFFA 866

QY 143 TMFPHVIA-DYRVLPIEVPLEDVNRNL---FHNNDNAPTM--RLVCMLYTPLRASGSSS 196
Db 867 EFYHIVSTDSQAEPIREYKLLSLTLTALQSIDSNSHMGKLSRIY-----916

QY 197 GTDFEVIAGRVLTCPSPDFSLFLVPPNVEQKTFPSVNPPLNTLSNRVPSILKSMV 256
Db 917 -----LSSARMVTAVPAVFSKLVTM-----LNASGSTHFTMRRLMA 954

QY 257 SRDHQWQV-FQNG-RVTLDG---QLQGTPTTSASQLCKIRGVSFHANGNGYNLTELD- 310
Db 955 IADEVEIAEVQLGVEDTVGHQDSLQNAVPTSCLENSLSLEHTVHREKTKGLSATLSA 1014

QY 311 -----GSPYHAFESPAPIGFPDGLGCDWHMEAS 338
Db 1015 SSEDISDRLAGVSLPSTTTEQPKPAVQTKGRPHSQCLNSPLSHAQL-----MPPA 1068

QY 339 PTTQFNTGDVTKQINVKOESAFAP 362
Db 1069 PSAPCSAPSVPDISKRRPQAFVP 1092

RESULT 11
US-09-091-219-24
; Sequence 24, Application US/09091219
; Patent No. 6171592
; GENERAL INFORMATION:
; APPLICANT: STUDDERT, Michael J.
; APPLICANT: CRABB, Brendan S.
; APPLICANT: FENG, Li
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
; FILE REFERENCE: 040268/0151
; CURRENT APPLICATION NUMBER: US/09/091,219
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: PCT/AU96/00815
; EARLIER FILING DATE: 1996-12-18
; EARLIER APPLICATION NUMBER: AU PN7201
; EARLIER FILING DATE: 1995-12-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2318
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
; US-09-091-219-24

Query Match          3.7%; Score 106; DB 4; Length 2318;
Best Local Similarity 22.3%; Pred. No. 0.64;

Matches 119; Conservative 65; Mismatches 241; Indels 145; Gaps 33:

QY 9 PTNMDGTSGAGOLVPEANTAEPISEPVAGAAAT-AAATAGOVNMIDPWIMNNYVOAPOQE 67
Db 633 PVRMD---NAVPIVQPPAAQPLQIQ--SGVLTQGSCTPLMVATLHPQVA-----677

QY 68 FTISPNNTPGDIILFDLQLGPHLNPFLSHLAQMYNGWGNMKVVKVLLAGNAFTAGKIIISC 127
Db 678 -TITPQVA---VPFTLSCAAGRPAIVEQTAAVLQAMPGGTQ-QILLPS---TWOQL----725

QY 128 IPPGFAAQNTISIAQATWFPHVIAADVRLVLEPIEVLDEDVNRVLFNHNDNAPTMRLVCMLYT 187
Db 726 --PCVALHN-SVQPTAMPEAMSGQ-----QLADWRNAHSHGNQYSTIMQQQSLLTN 775

QY 188 -----PLRAS-----GSSSGTDPPEVIAGRV--LTCPSP 213
Db 776 HVTLATAQPLNVGVAHVVRQOQSSSLPSKKNKOSAPVSSKSLD--VLPSQVSVLWGSSP 833

QY 214 ---DFSEFLFLVPPNVEQKTKPFSPVNPPLNTLS-----NSRVPSLKSMMVYS 257
Db 834 LRTTSSYNSLVP--VQDQHQPITIPDTPSPVSVITIRSDTDEEDNKYKPS--SSGLKP 889

QY 258 RDHGMVQFQNGRVTLDGQLQGTPTTSASOLCKIRGVSFHFANGN---GYNLTELDGSPY 314
Db 890 RSN--VISYTVNDSPDSLSLSPYSTDTLSALRG-----NSSVLEGGPRVVA DGTGT 942

QY 315 HAFESPAPIGFPDGLGCDWHMEASPTTQFNTGDVTKQINVKOESAFAPHLGCTQAQGLS- 373
```

Db 943 RTIIVP-PLK-TQLGDCVTATQASGLLSNKTKPVASVSGSSGCCITPTCYRAQRGGTSA 1000
QY 374 -----DVSNTNMIKLGWSPVSDGHRGDVDPWVPIRYGSLTLEAAQLAPPIYPGFGGEA 429
Db 1001 AQLNLSONQOOSA-----APTQSRSSNPAP-----RRQAFVAP-----LSQA 1040
QY 430 IVFEMDSDFPTAHTNG-----LSVPCTTPOEFVTHFVNEQAPTRGEA-----ALLHYLDPD- 480
Db 1041 PYTFQHSPL-HSTGHPHLAPAPAHLPQ--AHLYTYAAPTSAALGSTSSIAHLFSPQG 1097
QY 481 THNLGEFKLYPEGEMTCVPNSSGTGPOTL 510
Db 1098 SSRHAAAYTHPSTLVHQVPVS--VGPSLL 1125
RESULT 13
US-08-460-860-4
; Sequence 4, Application US/08460860
; Patent No. 5665584
; GENERAL INFORMATION:
; APPLICANT: HATAMOTO, OSAMU
; APPLICANT: WATARAI, TERUO
; APPLICANT: MIZUSAWA, KIYOSHI
; APPLICANT: NAKANO, EIICHI
; TITLE OF INVENTION: A DNA FRAGMENT CONTAINING A TANNASE
; TITLE OF INVENTION: GENE, A RECOMBINANT PLASMID, A PROCESS FOR PRODUCING
; TITLE OF INVENTION: TANNASE, AND A PROMOTER
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUBSTADT
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,860
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 159973/1994
; FILING DATE: 12-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 83973/1995
; FILING DATE: 10-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REFERENCE/DOCKET NUMBER: 7127-001-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 588 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-460-860-4
Query Match 3.5%; Score 102.5; DB 1; Length 588;
Best Local Similarity 21.7%; Pred. No. 0.15;
Matches 66; Conservative 41; Mismatches 100; Indels 97; Gaps 18;
QY 273 LDQLOGTTPTTSASOLCKIRGSVFHANGNGYNLTGDSPYH-AFESPAPIGFPDLGEC 331
Db 264 LDGRTDGV--VSRTDLCKL-----NFNLTSIIIGEPYYCAAGTSTSLGF----- 304

QY 332 DHMEASPTTQFNTGVIKQINVKQE-----SAFAP-HLGTIQADGLSDVSV-----N 378
Db 305 -----GFSNG---KRSNVKROAEGSTTSYQPAQNCVTVTARGVAVAAIYDGLHNS 351
QY 379 TNMIAKLGW--VSPVSDG---HRGDVDPVW--IPRYGS-----TLTEAAQLAPPI 421
Db 352 KGERAYLSQJIASLSDAETEYNSDTGKWLNPSTCGEYVTKFIQLLNJNLSLNNVT 411
QY 422 YPGFGEAIVFEMSDFPPIAHGTNGLSVPCCTIPO-----EFVTHFVNQAPTRGEAAL 473
Db 412 Y-----DTLVDMNMTGNVRYMD---SLQTTLPDLTTFQSSGGKLLHYHGSDPSIPAASS 463
QY 474 LHYLDP-----DTHRNLFGEFKLYPEGFMTCPVPNSSGTGPOTLPINGVVF 518
Db 464 VHYQAVRSVMYGDKTEEEALEADWYQFYLLIP-GAAHCGTNSLQPGP--YPENNMEIM 520
QY 519 VSWV 522
Db 521 IDWV 524
RESULT 14
US-09-171-699-10
; Sequence 10, Application US/09171699
; Patent No. 6448389
; GENERAL INFORMATION:
; APPLICANT: The Wistar Institute of, Anatomy & Biology
; Gonczol, Eva
; Berencsi, Klara
; Kari, Csaba
; TITLE OF INVENTION: No. 6448389el Cytomegalovirus DNA Constructs and
; Uses Therefor
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Center, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,699
; FILING DATE: 19-Jan-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/015,717
; FILING DATE: 23-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kodroff, Cathy A.
; REGISTRATION NUMBER: 33,980
; REFERENCE/DOCKET NUMBER: WST66APCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1048 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-171-699-10
Query Match 3.5%; Score 102; DB 4; Length 1048;
Best Local Similarity 21.8%; Pred. No. 0.44;
Matches 120; Conservative 50; Mismatches 214; Indels 166; Gaps 25;

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	106	3.7	1302	12	US-10-000-864-2		Sequence 2, Appli
2	106	3.7	1493	10	US-09-858-754-3		Sequence 3, Appli
3	106	3.7	1493	12	US-10-000-864-8		Sequence 8, Appli
4	103.5	3.6	1170	12	US-10-135-687-2		Sequence 2, Appli
5	103.5	3.6	1210	10	US-09-860-352A-2		Sequence 2, Appli
6	100.5	3.5	1070	10	US-09-735-367B-6		Sequence 6, Appli
7	100	3.5	828	9	US-09-738-626-5038		Sequence 5038, Ap
8	100	3.5	1601	10	US-09-862-027-40		Sequence 40, Appl
9	100	3.5	2005	10	US-09-735-367B-3		Sequence 3, Appli
10	100	3.5	2063	10	US-09-735-367B-2		Sequence 2, Appli
11	98	3.4	1209	12	US-10-135-687-4		Sequence 4, Appli
12	98	3.4	1493	10	US-09-858-754-4		Sequence 4, Appli
13	97	3.3	956	9	US-10-121-032-63		Sequence 63, Appl
14	96	3.3	1344	9	US-09-738-626-6888		Sequence 6888, Ap
15	95.5	3.3	678	9	US-09-712-363-158		Sequence 158, App
16	95	3.3	504	9	US-09-738-626-3680		Sequence 3680, Ap
17	94.5	3.3	662	10	US-09-924-358-11		Sequence 11, Appl
18	94	3.2	1477	9	US-10-092-880-4		Sequence 4, Appli
19	93.5	3.2	498	10	US-09-925-297-673		Sequence 673, App


```
US-10-135-687-2
; Sequence 2, Application US/10135687
; Patent No. US20020123120A1
; GENERAL INFORMATION:
; APPLICANT: CHANDRAMOULISWARAN, Ishwar et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001068DIV
; CURRENT APPLICATION NUMBER: US/10/135,687
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 09/749,588
; PRIOR FILING DATE: 2000-12-28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-135-687-2

Query Match      3.6%; Score 103.5; DB 12; Length 1170;
Best Local Similarity 20.9%; Pred. No. 1.1;
Matches 119; Conservative 65; Mismatches 241; Indels 145; Gaps 33;

QY 9 PTNMDCTGAGQLVPEANTAEPISEMPVAGAAT-AAATAGQVNMIDPWIMNNYVQAPOGE 67
Db 633 PVKMD---NAVPIVQAPAAQPLQIQ--SGVLTQGGCTPLMWATLHPQVA----- 677

QY 68 FTISPNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNKVKVLLAGNAFTAGKIIISC 127
Db 678 -TITPOYA---VPFTLSCAAGRPALEQTAAVLQANPGGTQ-QILLPS---TWQOL---- 725

QY 128 IPGFAAQNISTAQATMFPHVADRVLEPIEVPLEDVNRVLFHNNNDNAPTMRVLCMLYT 187
Db 726 --PGVALHN-SVQPTAMIEPMGSGQ-----QLADWRNAHSHGNYSTIMOQPSLLTN 775

QY 188 -----PLRAS-----GSSSGTDPFVIAGRV--LTCPSP 213
Db 776 HVTATAQPLNVGVAVHVRQQSSSLPSKKNQKQAPVSSKSSLD--VLPQSVYSLVGSSP 833

QY 214 ---DFSFLFPLVPNEQKTKPFSVPLNPLNTLS-----NSRVPSLIKSMWVS 257
Db -834 LRTTSYNSLVP--VDQHQPIIIPDTPSPVSVITIRSDTDEEDNKYKPS--SSGLKP 889

QY 258 RHGQMVQFQNGRVTLDGQLQGTTPTSASQLCKIRGSVFHANGN---GYNLTLDGSPY 314
Db 890 RSN--VISYTVNDSPDSSLSSTDTLSALRG-----NSGVLEGPGRVVADGTGT 942

QY 315 HAFESAPIGFDLGCEDHMEASPTQFNTGDVIKQINVKQESAFAPHLGTIQADGLS- 373
Db 943 RTIIVP-PLK-TQLGCTVATQASGLLSNKTKPVASVSGSSGCCITPTGYRAQRGTS 1000

QY 374 ----DVSVTNMIKLGWVSPVSDGHRGDVDPWIPRYGSTLTTEAAQLAPPIYPGFG 429
Db 1001 AQPPLNSQNOQSSA-----APTSQERSNNAP-----RRQAFVAP-----LSQA 1040

QY 430 IVFFMSDFPIAHGTNG---LSVPCTIPQEFVTHFVNEQAPTRGEA-----ALLHYLD 480
Db 1041 PYTFQHGSP-LHSTGHPLAPAPALPSQ--AHLTYAAPTSAALGSTSTIAHLFSP 1097

QY 481 THRNLEGEKLYPEGMTCVPSNCGTGPQL 510
Db 1098 SSRHAAAYTHPSTLVHQPVVS--VGPSLL 1125

RESULT 5
US-09-860-352A-2
; Sequence 2, Application US/09860352A
; Patent No. US20020132785A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Curtis, Rory
```

```
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: 13305 NOVEL PROTEIN KINASE MOLECULES AND
; FILE REFERENCE: 38155-20016.00
; CURRENT APPLICATION NUMBER: US/09/860,352A
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/205,301
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1210
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-860-352A-2

Query Match      3.6%; Score 103.5; DB 10; Length 1210;
Best Local Similarity 20.9%; Pred. No. 1.2;
Matches 119; Conservative 65; Mismatches 241; Indels 145; Gaps 33;

QY 9 PTNMDCTGAGQLVPEANTAEPISEMPVAGAAT-AAATAGQVNMIDPWIMNNYVQAPOGE 67
Db 673 PVKMD---NAVPIVQAPAAQPLQIQ--SGVLTQGGCTPLMWATLHPQVA----- 717

QY 68 FTISPNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNKVKVLLAGNAFTAGKIIISC 127
Db 718 -TITPOYA---VPFTLSCAAGRPALEQTAAVLQANPGGTQ-QILLPS---TWQOL---- 765

QY 128 IPGFAAQNISTAQATMFPHVADRVLEPIEVPLEDVNRVLFHNNNDNAPTMRVLCMLYT 187
Db 766 --PGVALHN-SVQPTAMIEPMGSGQ-----QLADWRNAHSHGNYSTIMOQPSLLTN 815

QY 188 -----PLRAS-----GSSSGTDPFVIAGRV--LTCPSP 213
Db 816 HVTATAQPLNVGVAVHVRQQSSSLPSKKNQKQAPVSSKSSLD--VLPQSVYSLVGSSP 873

QY 214 ---DFSFLFPLVPNEQKTKPFSVPLNPLNTLS-----NSRVPSLIKSMWVS 257
Db 874 LRTTSYNSLVP--VDQHQPIIIPDTPSPVSVITIRSDTDEEDNKYKPS--SSGLKP 929

QY 258 RHGQMVQFQNGRVTLDGQLQGTTPTSASQLCKIRGSVFHANGN---GYNLTLDGSPY 314
Db 930 RSN--VISYTVNDSPDSSLSSTDTLSALRG-----NSGVLEGPGRVVADGTGT 982

QY 315 HAFESAPIGFDLGCEDHMEASPTQFNTGDVIKQINVKQESAFAPHLGTIQADGLS- 373
Db 983 RTIIVP-PLK-TQLGCTVATQASGLLSNKTKPVASVSGSSGCCITPTGYRAQRGTS 1040

QY 374 ----DVSVTNMIKLGWVSPVSDGHRGDVDPWIPRYGSTLTTEAAQLAPPIYPGFG 429
Db 1041 AQPPLNSQNOQSSA-----APTSQERSNNAP-----RRQAFVAP-----LSQA 1080

QY 430 IVFFMSDFPIAHGTNG---LSVPCTIPQEFVTHFVNEQAPTRGEA-----ALLHYLD 480
Db 1081 PYTFQHGSP-LHSTGHPLAPAPALPSQ--AHLTYAAPTSAALGSTSTIAHLFSP 1137

QY 481 THRNLEGEKLYPEGMTCVPSNCGTGPQL 510
Db 1138 SSRHAAAYTHPSTLVHQPVVS--VGPSLL 1165

RESULT 6
US-09-735-367B-6
; Sequence 6, Application US/09735367B
; Patent No. US20020151477A1
; GENERAL INFORMATION:
; APPLICANT: Gustafsson, Jan-Ake
; APPLICANT: Caltra, Francoise
; APPLICANT: Antonsson, Per
; TITLE OF INVENTION: NUCLEAR RECEPTOR COACTIVATOR
; FILE REFERENCE: 102093-100
; CURRENT APPLICATION NUMBER: US/09/735,367B
; CURRENT FILING DATE: 2000-12-12
```

```
; PRIOR APPLICATION NUMBER: US 60/174,544
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1070
; TYPE: PRT
; ORGANISM: mammal
US-09-735-367B-6

Query Match          3.5%; Score 100.5; DB 10; Length 1070;
Best Local Similarity 18.8%; Pred. No. 1.8;
Matches 102; Conservative 56; Mismatches 188; Indels 197; Gaps 25;

QY 14 GTGAGQLVPEANTAEPISEPVAGATAAATAGQV-----NMIDPWI-----MNNY 60
DB 542 GNSGAPQLQANQVQH-----AGGAGGPPQONQOVSHGPPNMQPMSLGMTHGNNMQ 594
QY 61 VQAPQGEFTISPNTTGDILDLQLGP-----HLNPFSLHQAQYNGWGNMKYKV 111
DB 595 QAGTSGVPQVNLNMQG---PQOGPPSQLMGMHQOIIVPSQGMVQQ---OGTLNPNQPMI 649
QY 112 LLAGNAFTAGKIIISCIPPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDRVNLVLF 171
DB 650 LSRALMPQGMVNV---PP---SQNLGPPSPQM-----677
QY 172 NNDNAPTMLVCMLYTPLRASGSSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQ-KTK 230
DB 678 -----TPPKQLSQGQPMMAPHNQMM---GPOQVLLQONPMIEQIMTN 719
QY 231 PFSVNPILPLNTLSNRV---PSLIKSMVSRDHGMQVQFQNGRVTLTGDL----QGITPTS 284
DB 720 QMOGNKQOFNTQNSQVMPGPAQIMRGPTNMOGNMVQFTG---QMSGQMLPQGGPVNNS 776
QY 285 ASQLCKIRGSVFHANGGNGYNLTLDGSPYHAFESPAPIGFPPDLGECDDWHM---EASPTT 341
DB 777 PSQVWIGIQGVLRPPG-----PSP-----HMAQQHGD PAT 806
QY 342 QFNT-----GVVIOKINVKQESAFAPHILGTITQADGLS-----373
DB 807 TANNDVLSLQMMQDVSIQOTNMVPPHYQAMQGNASGNHFSGHGMSNAPFSGAPNGQM 866
QY 374 -----DYSVNTNMLAKLGWSPVSDGHRGDVDPWVPIPRYGSTLTLEAAQL-----417
DB 867 SCGQNGPFPVKNKDWLTSLPVLNL-LQSDISAGHFG-----VNNKQNTNANKPKKKKPP 920
QY 418 -----APTYPPGGEALVFFMSDPPIAHGNTNGLSVPCPTIQEFVTHFVNE-QA 465
DB 921 RKKKNSQDNLNTPDTRPAGLEEA-----DQPLPFGEGQISLDNSGPK--LPFESNRPPA 972
QY 466 PTR 468
DB 973 PSQ 975

RESULT 7
US-09-738-626-5038
; Sequence 5038, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OKHAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
```

```
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: Jp 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: Jp 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: Jp 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5038
; LENGTH: 828
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5038

Query Match          3.5%; Score 100; DB 9; Length 828;
Best Local Similarity 22.4%; Pred. No. 1.4;
Matches 106; Conservative 50; Mismatches 143; Indels 174; Gaps 28;

QY 111 VLLAGNAFTAGKIIISCIPPGFAAQNISIAQATMFPHVIADVRV-LEPIEVPLEDRVNL 169
DB 394 VEIAGGTVDAGRTLVDGP---AMQPTMKVTR--PSELAGVDYSAETVIARLEEVGTV 448
QY 170 FHND-----NAPTRM-----LVCMLYTPLRASGSSSGTDPEVI---AGRVLT-----209
DB 449 AVSGDTLEVTPPTWRGDLTMSADLVEEVLRLEGLEA---IPTIIPTAPAGRLGLTAQKRRR 506
QY 210 -----CSPDPDFSLFLVPPNV-----EOKTRPFSVNPILPL-----NTL 242
DB 507 AVGHALAVAGYAEIITPSP-----FMDEPFDVWGLAADERRKTVSVLN-PLEARNVL 559
QY 243 SNRPVPSLIKSM---VSRDHG-----QMVQFQNGRVTLTGLOGITPTS---S 286
DB 560 STSLLPMLDAVKRNVARGHNDPSLFLGQQVAFEHG-----SGVSPMPSVASRPEES 611
QY 287 QLCCKIRGSVFHANGGNGYNLTLDGSPYHA-----FESPAPIG---FPDLGECDD 333
DB 612 VVAEL-----VDSLNPQLHVATVGTGNTIEFEGPWGKGRAYTFADA-----652
QY 334 HMEASPTTQFNTQDVIKQINVKQESAFAPHILGTTOA---DGLSDVSVNTNMLAKLGWSP 390
DB 653 -IESARAVARAAGVTLELANA---DALPWHGRCALLIDG-----TP 691
QY 391 VSDGHRGDVDPWVPIRYG-----STLTAQAQLAPPYPPGGEALVFFMSDPP 438
DB 692 V--GYAGELHPQILEKAGLPARTCAMELDSLALPLVENLPAPV-----LSSEP 737
QY 439 IAHGTNGLSVPCPTIQEFVTHFVNEQAPTRGEAALLHYLDPPDTHRN--LGEPK 489
DB 738 ALHQDIALVDEITPAEDYRAVVEAGAGELIETVELF---DVRSEQLGENK 786

RESULT 8
US-09-862-027-40
; Sequence 40, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. US20020142428A1el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 1601
; TYPE: PRT
; ORGANISM: C. elegans
US-09-862-027-40

Query Match          3.5%; Score 100; DB 10; Length 1601;
```


Qy	172	NNDNAPTRLVCMILYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSFLEFLVPPNVQEQ-KTK	230
Db	620	-----TTPKQMLSQOGPQMMAHQNMM-----GPOGOVLLQONPWIEQIMTN	661
Qy	231	PFSVPNLPLNTLSNSRV---PSLTKSMWVRDGHQMVQFONGRVTLDGQL---QGTTPPTS	284
Db	662	QMOGNKQOFTNQNSNVMYPGPAQIMRGPTPNMOCNMVQFTG---QMSGOMLPQOGGPVNN	718
Qy	285	ASOLCKIRGSVFHANGNGYNLTLDGSPYHAFESPAPIGFDPOLGECDDWHM---EASPTT	341
Db	719	PSQVMYGOGVLRPPG-----PSP-----HMAOOHQDPAT	748
Qy	342	QFNT---GDVIKQINVKQBSAFAPHLTGTTQADGLS-----	373
Db	749	TANNDVLSLQMWMPDVSTQOTNMVPPHVOAQMGNSASGHSFGHGMSPFNAPFSGAPNCNM	808
Qy	374	-----DVSVNTNMIAKLGWSPVSDGHRGDVDPMVIPRYGSTLIEAAQL----	417
Db	809	SCGQNPFPVKNKDVLTSPLLVNL-LQSDISAGHFG-----VNNKQNTNANKPKKKPP	862
Qy	418	-----APPLYPPGGEAIVFEMSDPPTAHGTNGLSVPTIPO	454
Db	863	RKKNSOODLNTPTDRAGLEEA-----DOPPLPGEGISLDSGPK	904
RESULT 10			
US-09-735-367B-2			
; Sequence 2, Application US/09735367B			
; Patent No. US20020151477A1			
; GENERAL INFORMATION:			
; APPLICANT: Gustafsson, Jan-Ake			
; APPLICANT: Calra, Francoise			
; APPLICANT: Antonsson, Per			
; TITLE OF INVENTION: NUCLEAR RECEPTOR COACTIVATOR			
; FILE REFERENCE: 102093-100			
; CURRENT APPLICATION NUMBER: US/09/735,367B			
; CURRENT FILING DATE: 2000-12-12			
; PRIOR APPLICATION NUMBER: US 60/174,544			
; PRIOR FILING DATE: 2000-01-05			
; NUMBER OF SEQ ID NOS: 18			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 2			
; LENGTH: 2063			
; TYPE: PRT			
; ORGANISM: Human			
US-09-735-367B-2			
Query Match 3.5%; Score 100; DB 10; Length 2063;			
Best Local Similarity 18.6%; Pred. No. 5.3;			
Matches 98; Conservative 53; Mismatches 183; Indels 194; Gaps			
Qy	14	GTSAGAGQLVPEANTAEPISEMPVAGATAAATAGV-----NMIDPWI-----MNNY	60
Db	542	GNSGAPOLQANQNVQH-----AGGACAGPPQNMQVSHGPPNMMPQSLMGTHGNKNQ	594
Qy	61	VQAPQGEFTTSPNNTPGDILFDLQLGP-----IILNPFSLHAQMYNGWGNKKVKV	111
Db	595	QAGTSGVPQVNLNNQ-----QQQGPSPSLMGHQQIIVPSQGMVQO-QGTLPNPNP	649
Qy	112	LLAGNAFTAGKIIISCTPPGFAAQNISIAOATMPHVIADRVLEPIEVPLEDVNRVLFH	171
Db	650	LSRAQLMPQGMVYN--PP---SQNLGSPQRM-----	677
Qy	172	NNDNAPTRLVCMILYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSFLEFLVPPNVQEQ-KTK	230
Db	678	-----TTPKQMLSQOGPQMMAHQNMM-----GPOGOVLLQONPWIEQIMTN	719
Qy	231	PFSVPNLPLNTLSNSRV---PSLTKSMWVRDGHQMVQFONGRVTLDGQL---QGTTPPTS	284
Db	720	QMOGNKQOFTNQNSNVMYPGPAQIMRGPTPNMOCNMVQFTG---QMSGOMLPQOGGPVNN	776
Qy	285	ASOLCKIRGSVFHANGNGYNLTLDGSPYHAFESPAPIGFDPOLGECDDWHM---EASPTT	341

Db 777 PSQVMIQGVLRPPG-----PSP-----HMAQQHGDPA 806
Qy 342 QFNT---GDVTKQINVKQESAFAPHLGTIQADGLS----- 373
Db 807 TANNDVLSQMMPDVSIQDTNMVPPHVMQAMQNSASGNHFSHGMSFNAPFSGAPNGQM 866
Qy 374 -----DVSNTNNIAKLGVSPVSDGHRGDVDPWIPRYGSTLTLEAAQL----- 417
Db 867 SCQNGFPPVKNKDVTLTSLVNL-LQDISAGHEG-----VNNKQNTNANKPKKKPP 920
Qy 418 -----APPYIPGGEAIVFMSDFPIAHGTNGLSVPCITPQ 454
Db 921 RKKNSQDLNTPDTRPAGLEEA-----DQPLPGEQISLDSGPK 962

RESULT 11
US-10-135-687-4
; Sequence 4, Application US/10135687
; Patent No. US20020123120A1
; GENERAL INFORMATION:
; APPLICANT: CHANDRAMOULISWARAN, Ishwar et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: THEREOF
; CURRENT FILING DATE: 2002-05-01
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1209
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-135-687-4

Query Match 3.4%; Score 98; DB 12; Length 1209;
Best Local Similarity 20.9%; Pred. No. 3.6;
Matches 119; Conservative 67; Mismatches 239; Indels 144; Gaps 32;

Qy 9 PTNMDGTSGAQLVPEANTAEPISEMPVAGAAAT-AAATAGOVNMIDPWIMNMYVQAPQ-- 65
Db 673 PVRMD---NAVPIVQAPAAQLQIQ--SGVLTOGSCPTLMVATLHPQVA---TITQYA 724
Qy 66 GFTTISPNNTRGDIILDLQLGPHLNPFLSHLAQMYNGWGNKVKVLLAGNAFTAGKIII 125
Db 725 VPFTLSCAGRPA-----LVEQTAAVLQAMPGGTQ-QILLP-----SAWQOL- 764
Qy 126 SCIPPGFAAQNLSIAQATMFHVIADRVLEPIEVPLEDVNVLFNHNDNAPTMRVLCML 185
Db 765 ----PCVALHN-SVQPAAVIPEAMGSSQ-----QLADWRNNAISHGNOYSTIMQOQSLL 812
Qy 186 YT-----PL-----RAGSSS-----GTDPFVIAGRVLTCPSPDFSFL----- 218
Db 813 THNVTLTAQPLNVLGVVHVRRQQSSSLPSKKNKQAPVSSKSSLEVLPQSQVYSLVGSSP 872
Qy 219 -----FLVPPNVQKTKPFSVNLPLNTLS-----NSRVPSLIKSMVVS 257
Db 873 LRTTSSYNLVP--VQDQHQPIIIPDTSPVPVSTIIRSDTDEEDNKYEPI--SSSLKA 928
Qy 258 ROHGQWQFQNGRVTLGDLGQGTTPTSASOLCKIRGSVFHANGNGYNLTDLGSPYHAF 317
Db 929 RSN--VISVTYNDSPDSLSLSSPHSTDTLSALR-----GNSGTLLGPGRPAADG 978
Qy 318 ESPAPICFP---DLGECMDHMEASPTTQFNTGDKVTKQINVKQESAFAPHLGTIQADGLS 373
Db 979 IGRFTIIVPPLKTLQDCTVATQASGLLSKTKPVASVSGSSGCCITPTGYRAQRGGAS 1038
Qy 374 DV---SVNTNMIAKLGVSPVSDGHRGDVDPWIPRYGSTLTLEAAQAPPIYPPGFGAI 430
Db 1039 AVQPLNLSQNOQS-----SSASTSQBERSNP--APR-----RQQAQFVAP-----LSQAP 1080

Qy 431 VFMSDFPIAHGTNG---LSVPCTIPIQEFVTHFVNQCAPTRGEA-----ALLHYLDQD-T 481
Db 1081 YAFQHGSP-L-HSTGHPLAPAPAHLPSPQ---PHLYTYAAPTSAALGSTSIAHLFFPQGS 1137
Qy 482 HRNLGEKFLYPEGFMTCVPNNSGTGPOTL 510
Db 1138 SRHAAAYTHPSTILVHQVPVS--VGPSLL 1164

RESULT 12
US-09-858-754-4
; Sequence 4, Application US/09858754
; Patent No. US20020055130A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Gary L.
; TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING APOPTOSIS
; FILE REFERENCE: CPI-042
; CURRENT FILING DATE: 2001-05-16
; PRIOR FILING DATE: 1998-02-13
; PRIOR FILING DATE: 1997-02-14
; PRIOR FILING DATE: 1997-02-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1493
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-858-754-4

Query Match 3.4%; Score 98; DB 10; Length 1493;
Best Local Similarity 18.2%; Pred. No. 4.9;
Matches 83; Conservative 58; Mismatches 129; Indels 186; Gaps 22;

Qy 93 LSHLAQMYNGWGNMKV--KVLLAGNAFTAG-KIISCIPPGFAAQN----- 136
Db 694 ISTLLELCKGQAGELAVGREILKAGSIGVGVYVLCILGNQAGESNNWQELLRLCLID 753
Qy 137 --ISIAQATMFHVTIA-DVRVLEPIEVPLEDVNVL---FHNNDNAPT--RLVCMLYTP 188
Db 754 RLLETLSAEFYPHIVSTDVDSQAEPIEIRKLLSLLAFALQSIDNSHSMVGKLSRRIV-- 811
Qy 189 LRAGSSSGTDPFVIAGRVLTCPSPDFLFLVPPNVQKTKPFSVNPPLNTLSNRVP 248
Db 812 -----LSSARVMTVPPLFSKLVMT-----LSAGSSSHPA 841
Qy 249 SLIKSMVSRDH---QGVQFQNGRVTLDSQ---LQGTTPTSASOLCKIRGSVFHANGN 302
Db 842 RMRRLMATADEVEIAEVIQL-GSEDTLGGQDSSQALAPPYPRESSSLEHTAHVEKTK 900
Qy 303 GYNLTLDGSPYHAFESPA--PIGFPDLGECMDHMEASPTTQFNTGDKVTKQINVKQESAF 360
Db 901 GLKATRLSASSEDISDLAGVSVGLP-----SSATTE----- 932
Qy 361 APHLGTTQADGLSDSVSVNTNMIAKLGVSPVSDGHRGDVDPWIPRYGSTLTLEAAQLAPP 420
Db 933 -QPKPTVQTKG-----RPHSQCLNSSLPLSP 957
Qy 421 --IYPPGFGAIVFMSDFPIAHGTNGLSVPC---TIQEFVTHFVNQCAPTRGEAALL 474
Db 958 QLMFP-----AISAPCSSAFSVFAGSVT----- 980
Qy 475 HYLDPDTHRLNGEFLKYPEGFMTC-VPNNSGTGPOT 509
Db 981 ---DASKHR-----PRAFVPCIKIPAS---PQT 1002

RESULT 13
US-10-121-032-63
; Sequence 63, Application US/10121032
; Patent No. US20020155550A1

GENERAL INFORMATION:
APPLICANT: Bylina, Edward J.
TITLE OF INVENTION: GLYCOSIDASE ENZYMES
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gray Cary Ware & Freidenrich LLP
STREET: 4365 Executive Drive, Suite 1600
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/121,032
FILING DATE: 09-Apr-2002
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/09/134,078
FILING DATE: 13-AUG-1998
APPLICATION NUMBER: 08/949,026
FILING DATE: 10-OCT-1997
APPLICATION NUMBER: 60/056,916
FILING DATE: 06-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09010/024002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858/677-1456
TELEFAX: 858/677-1465
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 956 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 63:
US-10-121-032-63

Query Match 3.3%; Score 97; DB 9; Length 956;
Best Local Similarity 20.4%; Pred. No. 3.1;
Matches 137; Conservative 64; Mismatches 225; Indels 246; Gaps 34;

QY 5 SKDAPNM--DGTSGAGQLVPEANTAEPIIS-----MEPVAGAAATAAGQV 49
DB 191 SKD-PDNLIIIVGTSNYSQQVDVA-SADPISDTNVAYTLHFAAFNPHDNLNVAQTALDN 248
QY 50 NM---IDPMIMNYYVQAPQGETISPNTPFGDILLDLQLGPHLNPFLSHLAQMYNGWV-- 104
DB 249 NVALFVTEW--GTILNTGOGEPDKESTNTWMAFLKEKGIS-HANWSLSDKAPETGTSVYQ 305
QY 105 GNMKVKVLLAGNAFTAGKIIISCI-----PPGFAAQNISIAQAATMEPHVIADVRV 154
DB 306 AGQGVSGLSNKLTSAGEIVKNIQWDTETGTGKTTQCSITIECIRAMETAQAGDEII 365
QY 155 LEPIEVLPELDRVNLPHNNNDNAPTRLVCMLYPLRASGSSGTDPFVIAGRLVLTCPSPD 214
DB 366 IAPGNYNFQDKIQGAFNRS-----VY--LYGSANGSTNPILILRGESATNP-PV 411
QY 215 FSFL-----FLVP-----PNVEOKT-----KPFSPVNLPLNTL----- 242
DB 412 FSLGLDYNNGYLLSIEGDYWNIRKIDFKTSGKGIVLDSNNGSKLKLNVVHDIGEEAHLRD 471
QY 243 --SNSRV-----PSLIKSMWVRDHGMVQFONG-----RVTLDS 275
DB 472 GSNNSIDCCTTYNTRTPRGFGEGLYGSDKGQHDITYERACNNNTIENCTVGPNNVTAG 531
QY 276 --OQOCTTPTSASQLCKIRGSVFHANGNGYNLTE-----LDGSPYH 315

DB 532 VDVKECTMTI-----IRNCVFSABGISENSSDAFIDLKCAYGFEVYRNTFNVDGSEV- 584
QY 316 AFESPAPIGFDPDGECDWHMEASPTTQFNTGDVIKOINVKQESAFAPHLGTIOADGLSDV 375
DB 585 ----INTGVDFLDRG-----TGFNTG---FRNAIFENTY--NLGS----- 615
QY 376 SYNTNMIAKLGVSPVSDGHRGDVDPWVIPRYXGSTLUEAAQLAPPIYPPCFGEAIVFFMS 435
DB 616 -----RASEISTARKKOGSP-----EQTHVMDNIRNPN-----SV 645
QY 436 DFPPIAGHTNCL-----SVPCTIPQEFVHFVNEQAPTGEAALLHYLDP----- 479
DB 646 DFPISDGTENLVNKFPCPDWNIIEPCNPVDE-----TNOAPT-----ISFLSPVNNITLV 693
QY 480 -----DTHRNLGEEFLYPEGFTCVPNSS-----GTGPQTLPIGCVFVFSW 521
DB 694 EGYNLQVEYNATDADGTIDNVKLYIDNNLVRQINSTYSYKNGHSDSPNTDELNCL----- 747
QY 522 VSRFYQLKPVGT 533
DB 748 TECTYTLKAIAAT 759
RESULT 14
US-09-738-626-6888
Sequence 6888, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6888
LENGTH: 1344
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6888

Query Match 3.3%; Score 96; DB 9; Length 1344;
Best Local Similarity 18.4%; Pred. No. 6.3;
Matches 90; Conservative 57; Mismatches 167; Indels 176; Gaps 22;

QY 10 TNNMGTSAGQLVPEANTAEPIISMPEVAGAAATAAGVNMIDPMIMNYYVQAPQGET 69
DB 147 TNNMGSDGFEYVEITNTAEPIDFSYT-----LNYLY-PODEFT 185
QY 70 ISPNNT-----PGDIL-----FDLQLGPHLNPFLSHLAQMYNGWGNMKVKVLL 113
DB 186 ----NTNEAVAAEPDGVIIQPKSLVFWIKNGPNDEATAADNAEY----- 228
QY 114 AGNAFTAGKIIISCIPPGFA---AQNISTAQAATMFPHVI-----ADRVLEPIEV 160
DB 229 -GTNLEAGKDLVEISSGGMANGTARGMQTQNT--GHIVNRPYFNAGASDVKANEGLHP 285

Db 75 FYSNPSTGTFARAVKNNLFGG-----DLQGG-----STITQYVKNALVGSAGHQHSG 124
QY 106 NM-KVKVLLAGNAFTAG-----KIIISCIPGFAAQNISIAQATMPHPHVIADVRVLEP 157
Db 125 LMRKAKELVIATKMSGEWSKDDVLOQAYLNIYIFGRGAYGISAAKAYFDKPKVEQLTVAG- 183
QY 158 IEVPLEDVRNVLPHNNDNAPTMRVCMMLYTPLRASGSSSGTDP-----EVIAGRVL 209
Db 184 -----GALLAALI-----RPSLTDPAVDPEGAHARWNVLDGMVET 220
QY 210 ---CPSPDFSFLF--LVPPNV---EOKTRP-----FVSPNPLPLNTLSNS 245
Db 221 KALSPNDRAAQVFPETVPPDLARAENQTRGPNGLIERQVTRLELLEFNIDEQTLNT---- 276
QY 246 RVPSLIKSMVMVRDHGMVQFONGRVTLTGQ-----LOGTTTPTSASQLCKI-- 291
Db 277 -----QGLVVT-----TTIDPQAORAAEKAVAKYLDGQDPMRAAVSIDP 317
QY 292 -RGSVFEHANGG---NGYNLTETL---DGSYPYHAFESPAP-----IGFPDLGECDMHMEASPT 340
Db 318 HNGAVRAYYGGDNANGFDEAAGLQTSSEKFKFALVAALQEGIG---LG---YQVDSPL 371
QY 341 TOFNTGDIVIKOINVKQESAFAPHLGTTIQADGLSDVSVNTN-----MIAKLGWVSPVSD-GH 395
Db 372 TV-----DGIKITNVEGEC-----GTCNIAEALKMSLNTSYRLMLKNGGPOAVADAHAH 422
QY 396 RGDVDPWVPIPRYGSTLLEAAQLAPPIYPPGFGEAIVF-----FMSDFPIAHGT---NGLS 447
Db 423 QAGIAS-SFPGVAHTLSDGKGPP-----NNGIVLGQYQTRVIDMASAYATLAASGIY 475
QY 448 VPCTIPIQEFVTHFVNEQAPTRGEAALLHYLDPDTHRNLFGEFKL 490
Db 476 HP-----PHFVQKVVSANGV-----LFDASTADNTGDQRI 506

Search completed: January 16, 2003, 10:00:14
Job time : 22 secs

QY 161 PLED---VRNVLFHNNDNAPTMRVCMMLYTPLRASGSSSGTDPFVIAGRVLTC-PSPDFS 216
Db 286 AVDESLLKQILVLSGATPGTVVTSQIPNLSLAVIADSSV-PLITONTATSINPAEPFT 344
QY 217 FLFLVPPNVQOKTKFSPVN-----LPLNTLSNSRVPSLIKSM 255
Db 345 FAFNITDDVQVRTAFLHVTSSAGEAATTINLTEDDGSFNWALPAADLTGK---SWFETV 401
QY 256 VSRDHGMVQFONGRVTLTGQ-----LOGTT----- 281
Db 402 TATDGFNSVTTPEVRVTVDGANTDPLRLNLEBNQVSGTTDVGASDVFGDKLELLIDDA 461
QY 282 -----PTSASQLCKIRGSVFHANG--GNGYNLTETLSDSPY---HAFESPAPICF 325
Db 462 PAVTNSLSAAPTFAFMEVQT--DVFFRNGILAGGEELRIFDQGTYANTETISTPVPL-- 517
QY 326 PDLGECDDWHM-----EASPTQFNTGDIVIKOINVKQESAFAPHLGTRIADG 371
Db 518 -----YHINEDGLTVSVYAGTKAAPEIDLNNN--DDFOIRNLRLILPDGRTLITAG 568
QY 372 LSDSVSVNTNM 381
Db 569 ISDSNAWLNM 578

RESULT 15
US-09-712-363-158
; Sequence 158, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 158
; LENGTH: 678
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-158

Query Match 3.3%; Score 95.5; DB 9; Length 678;
Best Local Similarity 21.6%; Pred. No. 2.5; Indels 199; Gaps 35;
Matches 126; Conservative 63; Mismatches 195;
QY 7 DAPTNM-----DGTSGAGQLVPEANTAEPISMEPVAGATAAATAGQVNMIDFWIMNN 59
Db 24 DIRTNOVSTILASDGESEIAKIVPPEGNRVD-VNLSQVPMHVRAVIAAE-------DRN 74
QY 60 YVOAPOGEFT----ISPNNTPGDILFDLQGLPHLNFSLHAQMY-----NGWVG 105

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 09:54:55 ; Search time 147 Seconds
(without alignments)
2390.337 Million cell upd

Title: US-09-926-799-1

Perfect score: 2896

Sequence: 1 MMASKDAPTNDGTSGAGQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Learning rates: 0.00002
Gapop 10.0 , Gapext 0.5

Searched: 456914 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Fast processing: Minimum Match 0%
Maximum Match 100%

Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA Main: *

1:	/cgn2_6/pdodata/1/paa/US07PCTUS_COMB.pcp.*
2:	/cgn2_6/pdodata/1/paa/US07_COMB.pcp.*
3:	/cgn2_6/pdodata/1/paa/US06_COMB.pcp.*
4:	/cgn2_6/pdodata/1/paa/US080_COMB.pcp.*
5:	/cgn2_6/pdodata/1/paa/US081_COMB.pcp.*
6:	/cgn2_6/pdodata/1/paa/US082_COMB.pcp.*
7:	/cgn2_6/pdodata/1/paa/US083_COMB.pcp.*
8:	/cgn2_6/pdodata/1/paa/US084_COMB.pcp.*
9:	/cgn2_6/pdodata/1/paa/US085_COMB.pcp.*
10:	/cgn2_6/pdodata/1/paa/US086_COMB.pcp.*
11:	/cgn2_6/pdodata/1/paa/US087_COMB.pcp.*
12:	/cgn2_6/pdodata/1/paa/US088_COMB.pcp.*
13:	/cgn2_6/pdodata/1/paa/US089_COMB.pcp.*
14:	/cgn2_6/pdodata/1/paa/US090_COMB.pcp.*
15:	/cgn2_6/pdodata/1/paa/US091_COMB.pcp.*
16:	/cgn2_6/pdodata/1/paa/US092_COMB.pcp.*
17:	/cgn2_6/pdodata/1/paa/US093_COMB.pcp.*
18:	/cgn2_6/pdodata/1/paa/US094_COMB.pcp.*
19:	/cgn2_6/pdodata/1/paa/US095_COMB.pcp.*
20:	/cgn2_6/pdodata/1/paa/US096_COMB.pcp.*
21:	/cgn2_6/pdodata/1/paa/US097_COMB.pcp.*
22:	/cgn2_6/pdodata/1/paa/US098_COMB.pcp.*
23:	/cgn2_6/pdodata/1/paa/US099_COMB.pcp.*
24:	/cgn2_6/pdodata/1/paa/US100_COMB.pcp.*
25:	/cgn2_6/pdodata/1/paa/US101_COMB.pcp.*
26:	/cgn2_6/pdodata/1/paa/US102_COMB.pcp.*
27:	/cgn2_6/pdodata/1/paa/US050_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2896	100.0	545	23	US-09-926-799-1	Sequence 1, Appli
2	2605.5	90.0	544	6	US-08-273-257-8	Sequence 8, Appli
3	1379	68.3	544	23	US-09-926-799-4	Sequence 4, Appli
4	1977.5	68.3	546	23	US-09-926-799-3	Sequence 3, Appli
5	1950	67.3	530	3	US-07-696-454-3	Sequence 3, Appli
6	1950	67.3	530	3	US-07-941-365C-3	Sequence 3, Appli

Query Match 100.0%: Score 2896: DB 23: Length 545:

```

; REFERENCE/DOCKET NUMBER: DHHS Ref. No. E-163-94/0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-273-257-8

Query Match 90.0%; Score 2605.5; DB 6; Length 544;
Best Local Similarity 89.2%; Pred. No. 1e-241;
Matches 486; Conservative 24; Mismatches 34; Indels 1; Gaps 1;

Qy 1 MMASKDAPTNMDGTSAGQLVPEANTAEPIISMEPVAGAATAAATAGQVNMIDPWIMNNY 60
Db 1 MMASKDAPTNMDGTSAGQLVPEANTAEPIISMEPVAGAATAAATAGQVNMIDPWIMSNY 60
Qy 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVKVLLAGNAFTA 120
Db 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVKVLLAGNAFTA 120
Qy 121 GKIIISCIPPGFAAQNISIAQATMFPHVIADRVRLPIEVPLEDVRNVLPHNNDNAPTMR 180
Db 121 GKIIISCIPPGFAAQNISIAQATMFPHVIADRVRLPIEVPLEDVRNVLPHNNDSSPTMR 180
Qy 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVVEQTKPFSPVNPPLN 240
Db 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVVEQTKPFSPVNPPLN 240
Qy 241 TLSNRVPSLIKSMVMVRDHQMVQFQNGRVTLDGQLQGTTPTSASQLCKIRGSVFHANG 300
Db 241 TLSNRVPSLIKSMVMVRDHQMVQFQNGRVTLDGQLQGTTPTSASQLCKIRGSVFHANG 300
Qy 301 GNGYNLTLDGSPYHAFESPAPIGPDLGECDDHMEASPTTQFNTGDIKQINVKQESAF 360
Db 301 GNGYNLTLDGSPYHAFESPAPIGPDLGECDDHMEASPTTQFNTGDIKQINVKQESAF 360
Qy 361 APHLGTIQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWVPIRYGSTLTEAAQLAPP 420
Db 361 APHLGTIQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWVPIRYGSTLTEAAQLAPP 420
Qy 421 IYPPGGEAIVFFMSDFPIAHGTNGLSVPCCTIPEQEVTHFVNEQAPTRGEAALLHYLDPD 480
Db 421 IYPPGGEAIVFFMSDFPIAHGTNGLSVPCCTIPEQEVTHFVNEQAPTRGEAALLHYLDPD 480
Qy 481 THRNLEEFKLYPEGFMTCVPNSSGTPQTLPLNGVVFVSWVSRYQLKPVGTAGPACRL 540
Db 481 THRNLEEFKLYPEGFMTCVPNSSGTPQTLPLNGVVFVSWVSRYQLKPVGTAGPACRL 540
Qy 541 GIRRS 545
Db 541 GIRRS 544

RESULT 3
US-09-926-799-4
; Sequence 4, Application US/09926799
; GENERAL INFORMATION:
; APPLICANT: TAKEDA, NAKAZU
; APPLICANT: NATORI, KATSURO
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: KAMATA, KUNIO
; APPLICANT: SATO, TOSHINORI
; APPLICANT: SATO, SEIYA
; TITLE OF INVENTION: Detection kit for SRSV
; FILE REFERENCE: 217039USOXPT
; CURRENT APPLICATION NUMBER: US/09/926,799
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: JP 11175928
; PRIOR FILING DATE: 1999-06-22
```

```

; REFERENCE/DOCKET NUMBER: DHHS Ref. No. E-163-94/0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-273-257-8

Query Match 90.0%; Score 2605.5; DB 6; Length 544;
Best Local Similarity 89.2%; Pred. No. 1e-241;
Matches 486; Conservative 24; Mismatches 34; Indels 1; Gaps 1;

Qy 1 MMASKDAPTNMDGTSAGQLVPEANTAEPIISMEPVAGAATAAATAGQVNMIDPWIMNNY 60
Db 1 MMASKDAPTNMDGTSAGQLVPEANTAEPIISMEPVAGAATAAATAGQVNMIDPWIMSNY 60
Qy 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVKVLLAGNAFTA 120
Db 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVKVLLAGNAFTA 120
Qy 121 GKIIISCIPPGFAAQNISIAQATMFPHVIADRVRLPIEVPLEDVRNVLPHNNDNAPTMR 180
Db 121 GKIIISCIPPGFAAQNISIAQATMFPHVIADRVRLPIEVPLEDVRNVLPHNNDSSPTMR 180
Qy 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVVEQTKPFSPVNPPLN 240
Db 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVVEQTKPFSPVNPPLN 240
Qy 241 TLSNRVPSLIKSMVMVRDHQMVQFQNGRVTLDGQLQGTTPTSASQLCKIRGSVFHANG 300
Db 241 TLSNRVPSLIKSMVMVRDHQMVQFQNGRVTLDGQLQGTTPTSASQLCKIRGSVFHANG 300
Qy 301 GNGYNLTLDGSPYHAFESPAPIGPDLGECDDHMEASPTTQFNTGDIKQINVKQESAF 360
Db 301 GNGYNLTLDGSPYHAFESPAPIGPDLGECDDHMEASPTTQFNTGDIKQINVKQESAF 360
Qy 361 APHLGTIQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWVPIRYGSTLTEAAQLAPP 420
Db 361 APHLGTIQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWVPIRYGSTLTEAAQLAPP 420
Qy 421 IYPPGGEAIVFFMSDFPIAHGTNGLSVPCCTIPEQEVTHFVNEQAPTRGEAALLHYLDPD 480
Db 421 IYPPGGEAIVFFMSDFPIAHGTNGLSVPCCTIPEQEVTHFVNEQAPTRGEAALLHYLDPD 480
Qy 481 THRNLEEFKLYPEGFMTCVPNSSGTPQTLPLNGVVFVSWVSRYQLKPVGTAGPACRL 540
Db 481 THRNLEEFKLYPEGFMTCVPNSSGTPQTLPLNGVVFVSWVSRYQLKPVGTAGPACRL 540
Qy 541 GIRRS 545
Db 541 GIRRS 544

RESULT 3
US-09-926-799-4
; Sequence 4, Application US/09926799
; GENERAL INFORMATION:
; APPLICANT: TAKEDA, NAKAZU
; APPLICANT: NATORI, KATSURO
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: KAMATA, KUNIO
; APPLICANT: SATO, TOSHINORI
; APPLICANT: SATO, SEIYA
; TITLE OF INVENTION: Detection kit for SRSV
; FILE REFERENCE: 217039USOXPT
; CURRENT APPLICATION NUMBER: US/09/926,799
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: JP 11175928
; PRIOR FILING DATE: 1999-06-22
```

; PRIOR APPLICATION NUMBER: JP 11-175928
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 544
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE
US-09-926-799-4

Query Match 68.3%; Score 1979; DB 23; Length 544;
Best Local Similarity 66.8%; Pred. No. 3.3e-181;
Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;

Qy 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGOVNMIDPWIMNNY 60
Db 1 MMASKDAPTSADGATGAGQLVPEVNTADPIDPVAGSSTALATAGQVNLDPWLIINF 60

Qy 61 VOAPOGEFTISPNTPGDILFDLQLGPHLNPFLSHLAQMYNGVGMKVKVLLAGNAFTA 120
Db 61 VOAPOGEFTISPNTPGDILFDLQLGPHLNPFLSHLSQMYNGVMRVRVVLAGNAFTA 120

Qy 121 GKIIISCIPPGFAAQNISIAQATMPHVIADRVVLEPIEVLDPEDVRNVLFNHNDNAPTMR 180
Db 121 GXVIVCCVPPGQSSTLSAQATLFPVHVIADRVTLDPVEVLDPEDVRNVLHNNDQPTMR 180

Qy 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDSEFLFVPPNVEOKTKPESVPLNPLN 240
Db 181 LVCMLYTPLRGSGGTDSEFVAGRVLTCPGDFNLFVPPTEQKTRPFTVNPNIPLK 240

Qy 241 TLSNSRVPSLIKSMVSRDHGMQVQFQNGRVTLDGQLGQTTPTTSASQLCKIRGSVFHANG 300
Db 241 YLSNSRIPNPIEGMSLSPDQTNQVQFQNGRCITDQGLGTTPTVPSVQLCKFRGRI--TSG 298

Qy 301 GNGYNLTLDGSPYHAFESPAPIGPFDLGECDDWHMEAS--PTQNTGDVIVKQINVKOES 358
Db 299 QRVNLTLDELGSPFAFAAPAGPFDLGSCDWHIEMSKIPNSSTQNNPIVTVNSKPNQSQ 358

Qy 359 AFAPHLGTQADGLSDVSVNTNMIKLGWSPVSDGHRGCDVDPWIPRYGSTLTEAAQLA 418
Db 359 QVPHLSSTILD--ENVSSGGYIGTIOWTSPDSGGANTFWKIPDYVGSGLAASQLA 416

Qy 419 PPIYPGGEALVFFMSDFPIAHGTNGLS----VPCITPQEEVTHFVNEQAPTRGEAALL 474
Db 417 PAVYPPGGENIVYFWASIP---GPNQSGSPNLVPCLLPQEVITHFISEQAPIQGEAALL 473

Qy 475 HYLDPDTHRNLEGEKLYPEGEFMTCPVNSSGTCGQPTLPINGVVFVSWVSRRFYQLKPVGTA 534
Db 474 HYVDPDTNRNLGEKLYPGGYLTCVPNSSSTGQQQLPLDGVFVFASWSRFFYQLKPVGTA 533

Qy 535 GPA-CRLGIRR 544
Db 534 GPARGLGVRR 544

RESULT 4
US-09-926-799-3
; Sequence 3, Application US/09926799
; GENERAL INFORMATION:
; APPLICANT: TAKEDA, NAOKAZU
; APPLICANT: NATORI, KATSURO
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: KAWATA, KUNIO
; APPLICANT: SATO, TOSHINORI
; TITLE OF INVENTION: Detection kit for SRSV
; FILE REFERENCE: 217039US0XPT
; CURRENT APPLICATION NUMBER: US/09/926,799
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: JP 11175928
; PRIOR FILING DATE: 1999-06-22

; PRIOR APPLICATION NUMBER: JP 11-175928
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 546
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE
US-09-926-799-3

Query Match 68.3%; Score 1977.5; DB 23; Length 546;
Best Local Similarity 66.6%; Pred. No. 4.6e-181;
Matches 367; Conservative 76; Mismatches 95; Indels 13; Gaps 7;

Qy 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGOVNMIDPWIMNNY 60
Db 1 MMASKDAPQSDAGSAGAGQLVPEVNTADPLMEPVAGPTTAVATAGOVNMIDPWIVNMF 60

Qy 61 VOAPOGEFTISPNTPGDILFDLQLGPHLNPFLSHLAQMYNGVGMKVKVLLAGNAFTA 120
Db 61 VQSPQGEFTISPNTPGDILFDLQLGPHLNPFLSHLSQMYNGVMRVRILLAGNAFSA 120

Qy 121 GKIIISCIPPGFAAQNISIAQATMPHVIADRVVLEPIEVLDPEDVRNVLFNHNDNAPTMR 180
Db 121 GXIIVCCVPPGFTSSSLTIAQATLFPVHVIADRVTLPIEPIEMPLEDVRNVLHNDNQPTMR 180

Qy 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDSEFLFVPPNVEOKTKPESVPLNPLN 240
Db 181 LVCMLYTPLRGSGGNSDSFVYVAGRVLTAPSDSFSLFVLPPTIEQKTRAFVTPNIPLQ 240

Qy 241 TLSNSRVPSLIKSMVSRDHGMQVQFQNGRVTLDGQLGQTTPTTSASQLCKIRGSVFHANG 300
Db 241 TLSNSRFPSLIOGMLSPDASQVQFQNGRCILIDQLLQGLTTTATSCQLFRVRGKI--NOC 298

Qy 301 GNGYNLTLDGSPYHAFESPAPIGPFDLGECDDWHMEASPT--TQFNTGDVIVKQINVKOE-S 358
Db 299 ARTLNLTEVDGKPFMAFSDPAPVGFDFGKCDWHMRISKTPNNTSSGDPMSRVSVQTNVO 358

Qy 359 AFAPHLGTQADGLSDVSVNTNMIKLGWSPVSDGHRGCDVDPWIPRYGSTLTEAAQLA 418
Db 359 GVPHLGSIQFDEVENHPTG--DYIGTIEWISQPSPTPGDIDNLWEIPDYVGSLSQAANLA 417

Qy 419 PPIYPGGEALVFFMSDFPIAHGTNGLS----VPCITPQEEVTHFVNEQAPTRGEAALL 474
Db 418 PVVFPFGGEALVYFVSAP---GPNRSAPNDVPCLLPQEVITHFVSEQAPTMGDAALL 474

Qy 475 HYLDPDTHRNLEGEKLYPEGEFMTCPVNSSGTCGQPTLPINGVVFVSWVSRRFYQLKPVGTA 534
Db 475 HYVDPDTNRNLGEKLYPGGYLTCVPNGVAGPQQQLPLNGVFLFVSWVSRRFYQLKPVGTA 534

Qy 535 GPA-CRLGIRR 544
Db 535 STARSRLGVRR 545

RESULT 5
US-07-696-454-3
; Sequence 3, Application US/07696454
; GENERAL INFORMATION:
; APPLICANT: Estes, Mary K.
; APPLICANT: Jiang, Xi
; APPLICANT: Graham, David Y.
; TITLE OF INVENTION: Methods and Reagents to Detect and
; TITLE OF INVENTION: Characterize Norwalk and Related Viruses.
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; STREET: Patent Department, Fulbright & Jaworski
; CITY: Houston
; STATE: Texas
; COUNTRY: USA

```

; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect converted to DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/696,454
; FILING DATE: 19910506
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene A.
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: D-5179CIP-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 651-5246
; TELEFAX: (713) 651-5246
; TELEX: Western Union 762829
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-696-454-3

Query Match 67.3%; Score 1950; DB 3; Length 530;
Best Local Similarity 67.0%; Pred. No. 2e-178;
Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;

QY 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNNY 60
Db 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSTAVATAGQVNPIDPWIIINF 60
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYKVLGNAFTA 120
Db 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYKVLGNAFTA 120
QY 121 GKIIISCIPIPGFAAONISIAQATMPEPHIVADRVLPVLEPIEVLDDVRLVFNHND-NAPTM 179
Db 121 GKIIISCIPIPGFASHNLITIAQATLPHVIAVRLDPIEVLDDVRLVFNHNDNQOTM 180
QY 180 RLVCMLYTLRASGSSSGTDPFVIAGRVLTCPSPDFSLFVPPNVEQKTFPSVPLNPL 239
Db 180 RLVCMLYTLPLRTGGTGC--DSFVAVGRVMTCPSPDFNLFVLPVTEQKTRFTLPLNPL 238
QY 240 NTLNSRVPSLIKSMVSRDHQGMQVQFNGRVTLTGQLOGTTTPTASQLCKIRGSVFHAN 299
Db 239 SSLNSRAPLPISSMGISPDNVQSVQFNGRCTLGRLVGTTPVSLSHVAKIRGT----S 294
QY 300 GGNGYNLTLDGSPYHAFESPAPIGFDPDLGECDDHMEASPTTQNTGDVTKQINVKQESA 359
Db 295 NGTVINLTLDGTPFHPFEGPAPIGFDPDLGECDDHIN---MTQGHSSQTYQYDVTTPDT 351
QY 360 FAPHLGTIQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWVPIRYGSTLTLEAAQLAP 419
Db 352 FVPHLGSIQANGIG----SGNVGVLSWISPPSPSGQVDLWKIPNKGSSITEATHLAP 407
QY 420 FIYPGFGCAIVFMSDPIAHNGNLISVPTCIQEFVTHEVNCQATRCGEAALLHYLDP 479
Db 408 SVYPPGFEVLVFFEMSKMP--GFCAYNLCLLPQEYIASHLASQAPTVGGAALLHYVDP 464
QY 480 DTHRNLGFEKLYPEGFMTCPVNSSGTGPQTLPINGVFVSVWSRFXQLKPVGTAGPA-C 538
Db 465 DTHRNLGFEKLYPEGFMTCPVNSSGTGPQTLPINGVFVSVWSRFXQLKPVGTAGPA-C 538
QY 539 RLGIIR 544
Db 525 RLGLRR 530

RESULT 6
US-07-941-365C-3
```

```

; Sequence 3. Application US/07941365C
; GENERAL INFORMATION:
; APPLICANT: Matson, David O
; APPLICANT: Estes, Mary K
; APPLICANT: Jiang, Xi
; APPLICANT: Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; TITLE OF INVENTION: Characterize Norwalk and Related Viruses
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski Patent Dept
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07/941,365C
; APPLICATION NUMBER: 19920908
; FILING DATE: 19920908
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene A.
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: D-5526
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-651-5246
; TELEFAX: 713-651-5246
; TELEX: Western Union 762829
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-941-365C-3
```

```

Query Match 67.3%; Score 1950; DB 3; Length 530;
Best Local Similarity 67.0%; Pred. No. 2e-178;
Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;

QY 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNNY 60
Db 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSTAVATAGQVNPIDPWIIINF 60
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYKVLGNAFTA 120
Db 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYKVLGNAFTA 120
QY 121 GKIIISCIPIPGFAAONISIAQATMPEPHIVADRVLPVLEPIEVLDDVRLVFNHND-NAPTM 179
Db 121 GKIIISCIPIPGFASHNLITIAQATLPHVIAVRLDPIEVLDDVRLVFNHNDNQOTM 180
QY 180 RLVCMLYTLRASGSSSGTDPFVIAGRVLTCPSPDFSLFVPPNVEQKTFPSVPLNPL 239
Db 181 RLVCMLYTLPLRTGGTGC--DSFVAVGRVMTCPSPDFNLFVLPVTEQKTRFTLPLNPL 238
QY 240 NTLNSRVPSLIKSMVSRDHQGMQVQFNGRVTLTGQLOGTTTPTASQLCKIRGSVFHAN 299
Db 239 SSLNSRAPLPISSMGISPDNVQSVQFNGRCTLGRLVGTTPVSLSHVAKIRGT----S 294
QY 300 GGNGYNLTLDGSPYHAFESPAPIGFDPDLGECDDHMEASPTTQNTGDVTKQINVKQESA 359
Db 295 NGTVINLTLDGTPFHPFEGPAPIGFDPDLGECDDHIN---MTQGHSSQTYQYDVTTPDT 351
QY 360 FAPHLGTIQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWVPIRYGSTLTLEAAQLAP 419
Db 352 FVPHLGSIQANGIG----SGNVGVLSWISPPSPSGQVDLWKIPNKGSSITEATHLAP 407
```



```

QY 420 PIYPGFGAIVFFMSDFPIAHGTNCLSVPCITPQEFVTHFVNEQAPTRGEAALLHYLDP 479
DB 408 SYVPGFGVLVFFMSKMP---GPGAYNLPCULLPQEIYISHLASEQAPTVGEAALLHYVDP 464
QY 480 DTHRLNGEFLKYPEGFMTCVPSNGSGTGPOTLPINGVVFVSVWSRYQLKPVGTAGPA-C 538
DB 465 DTGRNLGEFKAYPDGLTLCVPNGASSGPOQLPINGVVFVSVWSRYQLKPVGTASSARG 524
QY 539 RLGIIR 544
DB 525 RLGLRR 530

RESULT 7
US-07-941-365E-3
; Sequence 3, Application US/07941365E
; GENERAL INFORMATION:
; APPLICANT: Matson, David O
; APPLICANT: Estes, Mary K
; APPLICANT: Jiang, Xi
; APPLICANT: Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; TITLE OF INVENTION: Characterize Norwalk and Related Viruses
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski Patent Dept
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/941.365E
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene A
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: D-5526
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-651-3634
; TELEFAX: 713-651-5246
; TELEX: Western Union 762829
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-941-365E-3
Query Match 67.3%; Score 1950; DB 3; Length 530;
Best Local Similarity 67.0%; Pred. No. 2e-178;
Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;

QY 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGAATAAGQVNMIDPIMNNY 60
DB 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMPVAGSSTAVATAGQVNPIDPWIINF 60
QY 61 VOAPQGEFTISPNTPGDTLFDLQGLPHLNPFLSHLAQNYGWCWGMKVKVLLAGNAFTA 120
DB 61 VOAPQGEFTISPNTPGDTLFDLQGLPHLNPFLSHLAQNYGWCWGMKVKVLLAGNAFTA 120
QY 121 GKIIISCIIPPGFAAQNISIAQATMPHPHVIADVRVLEPTEVPLEDVRNVLFNHND-NAPTM 179
DB 121 GKIIISCIIPPGFGCSNLTIAQATLPHVIADVRVLEPTEVPLEDVRNVLFNHNDNQOTM 180
QY 180 RLVCMLYTLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEOKTPFSPVNLPL 239

```

```

DB 181 RLVCMLYTLRTGGTG--DSEFVAGRVMTCPSPDFNLFVLPVPTVEQTRPFTLPNLPL 238
QY 240 NTLNSRVPSLTKSNMVRSDHGMQVQFNGRVYTLDSQLOGLTTPTSASQLCKIRGVSVFHAN 239
DB 239 SLSLSNRAPLPTSSKMGISPDNVQSQVQFNGRCVTLDSGLRVLTTPVSLSHVAKIRGT---S 294
QY 300 GGNGYNLTDELDSGYHAFESPAPIGFPDILGECDMHMEASPTTQFNTGDYVIKQINVKOESA 359
DB 295 NGTVINLTDELDTGTPHPPEGPAPIGFPDILGECDMHIN---MTQFGHSSQTFQYDVTDPDT 351
QY 360 FAPHLGTIOADGLSDSVSYNTNMIAKLGWSPVSDGIRGVDVDPWIPRYGSTLTAAQLAP 419
DB 352 FVPHLGSIOANGIG---SCNYVGVLSWISPPSPHSGSQVDLWKIPNYGSSITEATHLAP 407
QY 420 PIYPGFGAIVFFMSDFPIAHGTNCLSVPCITPQEFVTHFVNEQAPTRGEAALLHYLDP 479
DB 408 SYVPGFGVLVFFMSKMP---GPGAYNLPCULLPQEIYISHLASEQAPTVGEAALLHYVDP 464
QY 480 DTHRLNGEFLKYPEGFMTCVPSNGSGTGPOTLPINGVVFVSVWSRYQLKPVGTAGPA-C 538
DB 465 DTGRNLGEFKAYPDGLTLCVPNGASSGPOQLPINGVVFVSVWSRYQLKPVGTASSARG 524
QY 539 RLGIIR 544
DB 525 RLGLRR 530

RESULT 8
US-08-386-365-3
; Sequence 3, Application US/08386365
; GENERAL INFORMATION:
; APPLICANT: Matson, David O
; APPLICANT: Estes, Mary K
; APPLICANT: Jiang, Xi
; APPLICANT: Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; TITLE OF INVENTION: Characterize Norwalk and Related Viruses
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski Patent Dept
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/386,365
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene A
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: D-5526
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-651-3634
; TELEFAX: 713-651-5246
; TELEX: Western Union 762829
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-386-365-3
Query Match 67.3%; Score 1950; DB 7; Length 530;
Best Local Similarity 67.0%; Pred. No. 2e-178;

```

Thu Jan 16 15:59:22 2003

Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;

QY 1 MMASKDAPNTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGOVNMDPWINNY 60
Db 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMPDVAGSSTAVATAGQVNPDPWINNF 60

QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYLLAGNAFTA 120
Db 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYLLAGNAFTA 120

QY 121 GKIIISCIPPGFAAQNSTIAQATMPHVIADRVLEPIEVLDPVNRVLFHND-NAPTM 179
Db 121 GKIIISCIPPGFAAQNSTIAQATMPHVIADRVLEPIEVLDPVNRVLFHND-NAPTM 179

QY 180 RLVCMLYTPLRASGSSGTPDPIAGRVLTCPSPDFSLFLVPPNVEQTKPESVNLPL 239
Db 180 RLVCMLYTPLRASGSSGTPDPIAGRVLTCPSPDFSLFLVPPNVEQTKPESVNLPL 239

QY 240 NTLNSRVPSLTKSMVSRDHGMVQFQNGRVTLTGQGLQGTTPTSASQLCKIRGSVFHAN 299
Db 240 NTLNSRVPSLTKSMVSRDHGMVQFQNGRVTLTGQGLQGTTPTSASQLCKIRGSVFHAN 299

QY 300 GNGYVNLTELDSGSPYHAFESPAPIGFDPDLGECGDHMEASPTTQFNTGDIKQINVKQESA 359
Db 300 GNGYVNLTELDSGSPYHAFESPAPIGFDPDLGECGDHMEASPTTQFNTGDIKQINVKQESA 359

QY 360 FAPHLGTIQADGLSDVSNTNMIKLGWSPVSDGHRGVDVWPVPRYGSTLTAEAAQLAP 419
Db 360 FAPHLGTIQADGLSDVSNTNMIKLGWSPVSDGHRGVDVWPVPRYGSTLTAEAAQLAP 419

QY 420 PIYPGFEAIFFMSDFPIAHGTNGLSVPCCTIPQEFVTHFVNEQAPTRGEAALLHYLDP 479
Db 420 PIYPGFEAIFFMSDFPIAHGTNGLSVPCCTIPQEFVTHFVNEQAPTRGEAALLHYLDP 479

QY 480 DTHRNLEGEFKLYPEGFMTCVPNSSGTPQLPINGVFFVSWSRFYQLKPVGTAGPA-C 538
Db 480 DTHRNLEGEFKLYPEGFMTCVPNSSGTPQLPINGVFFVSWSRFYQLKPVGTAGPA-C 538

QY 539 RLGIIR 544
Db 539 RLGIIR 544

QY 544 RLGIIR 544
Db 544 RLGIIR 544

RESULT 9
US-08-486-049-3
; Sequence 3, Application US/08486049
; GENERAL INFORMATION:
; APPLICANT: Estes, Mary K
; APPLICANT: Jiang, Xi
; APPLICANT: Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; TITLE OF INVENTION: Characterize Norwalk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE:
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-486-049-3

Query Match 67.3%; Score 1950; DB 8; Length 530;
Best Local Similarity 67.0%; Pred. No. 2e-178;
Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;

QY 1 MMASKDAPNTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGOVNMDPWINNY 60
Db 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMPDVAGSSTAVATAGQVNPDPWINNF 60

QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYLLAGNAFTA 120
Db 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVYLLAGNAFTA 120

QY 121 GKIIISCIPPGFAAQNSTIAQATMPHVIADRVLEPIEVLDPVNRVLFHND-NAPTM 179
Db 121 GKIIISCIPPGFAAQNSTIAQATMPHVIADRVLEPIEVLDPVNRVLFHND-NAPTM 179

QY 180 RLVCMLYTPLRASGSSGTPDPIAGRVLTCPSPDFSLFLVPPNVEQTKPESVNLPL 239
Db 180 RLVCMLYTPLRASGSSGTPDPIAGRVLTCPSPDFSLFLVPPNVEQTKPESVNLPL 239

QY 240 NTLNSRVPSLTKSMVSRDHGMVQFQNGRVTLTGQGLQGTTPTSASQLCKIRGSVFHAN 299
Db 240 NTLNSRVPSLTKSMVSRDHGMVQFQNGRVTLTGQGLQGTTPTSASQLCKIRGSVFHAN 299

QY 300 GNGYVNLTELDSGSPYHAFESPAPIGFDPDLGECGDHMEASPTTQFNTGDIKQINVKQESA 359
Db 300 GNGYVNLTELDSGSPYHAFESPAPIGFDPDLGECGDHMEASPTTQFNTGDIKQINVKQESA 359

QY 360 FAPHLGTIQADGLSDVSNTNMIKLGWSPVSDGHRGVDVWPVPRYGSTLTAEAAQLAP 419
Db 360 FAPHLGTIQADGLSDVSNTNMIKLGWSPVSDGHRGVDVWPVPRYGSTLTAEAAQLAP 419

QY 420 PIYPGFEAIFFMSDFPIAHGTNGLSVPCCTIPQEFVTHFVNEQAPTRGEAALLHYLDP 479
Db 420 PIYPGFEAIFFMSDFPIAHGTNGLSVPCCTIPQEFVTHFVNEQAPTRGEAALLHYLDP 479

QY 480 DTHRNLEGEFKLYPEGFMTCVPNSSGTPQLPINGVFFVSWSRFYQLKPVGTAGPA-C 538
Db 480 DTHRNLEGEFKLYPEGFMTCVPNSSGTPQLPINGVFFVSWSRFYQLKPVGTAGPA-C 538

QY 539 RLGIIR 544
Db 539 RLGIIR 544

QY 544 RLGIIR 544
Db 544 RLGIIR 544

RESULT 10
US-09-926-799-2
; Sequence 2, Application US/09926799
; GENERAL INFORMATION:
; APPLICANT: TAKEDA, NAOKAZU
; APPLICANT: NATORI, KATSURO
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: KAMATA, KUNIO
; APPLICANT: SATO, TOSHINORI
; APPLICANT: SATO, SEIYA
; TITLE OF INVENTION: Detection kit for SRSV
; FILE REFERENCE: 217039US0XPT
; CURRENT APPLICATION NUMBER: US/09/926,799
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: JP 11175928
; PRIOR FILING DATE: 1999-06-22

; PRIOR APPLICATION NUMBER: JP 11-175928
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 530
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE
US-09-926-799-2

Query Match 66.9%; Score 1938; DB 23; Length 530;
Best Local Similarity 67.0%; Pred. No. 2.9e-177;
Matches 366; Conservative 63; Mismatches 99; Indels 18; Gaps 7;

QY 1 MMASKADPTNMDGTSGAGOLVPEANTAEPISEMPVAGATAATAAGQVNMIDPWIMNNY 60
DB 1 MMASKADTSSVDGASGAGOLVPEVNASDPLANDPVAGSTAVATAGQVNPIDPWIIINF 60

QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGVGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGGVLFDSLGLPHLNPFLSHLSQMYNGVGMVNRVIRIMLAGNAFTA 120

QY 121 GKIIISCIPPGFAAQNISIAQATMPHVIADRVLEPIEVLPEIDVRNLFHND-NAPTM 179
DB 121 GKIIISCIPPGFSGHNLTAQATLPHVIAADVRLDPIEVLPEIDVRNLFHNNDRNQTM 180

QY 180 RLVCMLYTLPLRASGSSGTDPPFVIAGRVLTCPSPDFSEFLFVPPNVEOKTKPFSVNLPL 239
DB 181 RLVCMLYTLPLRTGGGTG--DSFVAGRVMTCPSPDFNFLLFVPPVVEQKTRPTLPNLPL 238

QY 240 NTLNSRVPSLIKSMVSRDHGQVQFNGRVTLDGLOLQGTTPTSASOLCKIRGSVFHAN 299
DB 239 SSLSNSRAPLPSIGMGISPDNVQSVQFNGRCTLDGRLVGTTVPVSLSHVAKIRGT---S 294

QY 300 GNGYNTLDELSPYHAFESPAPIGPPDILGECDMHMEASPTQFNTGQVVIKQINVKQESA 359
DB 295 NGTVINLTLDGTTFHFFEGPAPIGPPDLGGGDWHIN---MTQFGHSSQTYDVTDPDT 351

QY 360 FAPHILGTIOADGLSDVSNTNIAKLGWSPVSDGHRGVDVDPKVIIPRYGTLTEAAQLAP 419
DB 352 FVPHLGSIOANGIG----SGNTIGVLSWSPSPSHSGSGQVDLWKIPNYGSSITEATHLAP 407

QY 420 PIYPGFGGAIVFEMSDFFIAHGTNGLSVPCITPOEFVTHFVNEQAPTGEAALLHYLDP 479
DB 408 SVYPPGGEVLVFFEMSKIP--GPGAYSLPCLLPQEIYISHLASEQAPTYGEAALLHYVDP 464

QY 480 DTHRNILGEFKLPEGPMTCVPNSSGTPOTLPINGVFEVSVWSRYQLKPVGTAGPA-C 538
DB 465 DTGRTLGEFKAYPDGFLTCVPNGASSGPOQLPINGVFEVSVWSRYQLKPVGTASSARG 524

QY 539 RLGIIR 544
DB 525 RLGLRR 530

RESULT 11
US-09-926-799-7
; Sequence 7, Application US/09926799
; GENERAL INFORMATION:
; APPLICANT: TAKEDA, NAOKAZU
; APPLICANT: NATORI, KATSURO
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: KAMATA, KUNIO
; APPLICANT: SATO, TOSHINORI
; APPLICANT: SATO, SEIYA
; TITLE OF INVENTION: Detection Kit for SRV
; FILE REFERENCE: 217039USOXPCT
; CURRENT APPLICATION NUMBER: US/09/926,799
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: JP 11175928
; PRIOR FILING DATE: 1999-06-22

; PRIOR APPLICATION NUMBER: JP 11-175928
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 540
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE
US-09-926-799-7

Query Match 42.3%; Score 1225; DB 23; Length 540;
Best Local Similarity 46.7%; Pred. No. 2.2e-108;
Matches 261; Conservative 89; Mismatches 173; Indels 36; Gaps 14;

QY 1 MMASKADPTNMDGTSGAGOLVPEANTAEPISEMPVAGATAATAAGQVNMIDPWIMNNY 60
DB 1 MMASKADATPSNDGAAG---LVPESNN-EAMALEPVVGASLAAPVTGQTNIIDPWIRTNF 56

QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGVGMKVKVLLAGNAFTA 120
DB 57 VQAPNGEFTVSPNSPCEILVNLELOPELNPYLALHARMYNGYAGOMEVQVMLAGNAFTA 116

QY 121 GKIIISCIPPGFAAQNISIAQATMPHVIADRVLEPIEVLPEIDVRNLFH-NNDNAPTM 179
DB 117 GKIIIFAAVPPYFPVENLSPSQITMTMPHVIDVRLPEVLLPMPDVRSTLFHFQKDEPKM 176

QY 180 RLVCMLYTLPLRASGSSGTDPPFVIAGRVLTCPSPDFSEFLFVPPNVEOKTKPFSVNLPL 239
DB 177 RLVCMLYTLPLRSNG--SGDDVFTVSCRILTRPSPEEDFTYLVPPVVEKTKPTLPVLTL 234

QY 240 NTLNSRVPSLIKSMVSRDHGQVQFNGRVTLDGLOLQGTTPTSASOLCKIRGSVFHAN 299
DB 235 GELNSRFPPLSDIEWTSPNESIVQVQNGRVTLDELLGTLQLOACNCSIRGKVTGQV 294

QY 300 GNGY----NLTELDGSPYHAFES-PAPIGFPDL-GECDMH-----EASPTTQFNTGD 347
DB 295 PSEQHMMNLITNLNGTQDPTDDVPALGVDPDFAGEVFGVLSQRNKGESNPANRAHDAV 354

QY 348 VIKQINVKQESAFAPHLGTIOAD--GLSDVSNTNIAKLGWSPVSDGHRGVDVDPKVIIP 405
DB 355 V-----ATVSDYTKPLGLVQIGTWNNTDVENOQTKFTPIG-LNEVANGHR--FEQWTL 406

QY 406 RYGSTLTEAAQLAPPIYPGFGGAIVFEMSDFFIAHGTNGLSVPCITPOEFVTHFVNEQA 465
DB 407 RYSGALTNNLAPAVAPLPPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWQHFQESA 466

QY 466 PTRGEAALLHYLDPDTHRNILGEFKLPEGPMTCVPNSSGTPOTLPINGVFEVSVWSRF 525
DB 467 PSLGDVALVRYVNPDTGRVLFELAKLHKGGLTV--SSTSTGPPVVVPANGYFKRFDNVNQF 524

QY 526 YQLKPVGTAGPACRLGIIR 544
DB 525 YSLAPMGTGN-----GRRR 538

RESULT 12
US-08-273-257-5
; Sequence 5, Application US/08273257
; GENERAL INFORMATION:
; APPLICANT: LEW, Judy F.
; APPLICANT: GREEN, Kim Y.
; APPLICANT: VALDESUSO, Jose
; TITLE OF INVENTION: Calicivirus capsid genes and their uses
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,257
; FILING DATE: 11-JUL-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 15280-209
; REFERENCE/DOCKET NUMBER: DHHS Ref. No. E-163-94/0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-273-257-5

Query Match 40.9%; Score 1184.5; DB 6; Length 548;
Best Local Similarity 46.4%; Pred. No. 1.8e-104;
Matches 262; Conservative 79; Mismatches 181; Indels 43; Gaps 15;

QY 1 MMASKD-APTNDGTSAGQLVPEANTAEPISEMPVAGAAATAAGVNMDIPMNN 59
Db 1 MKMASNDAAPSN----DGAACLVPEINN--EAMALEPVAGAAIAAPLTGQNIIDPIMNN 55

QY 60 YVQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSLHAQMYNGWGNMKVKVLLAGNAFT 119
Db 56 FVQAPGGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFEVQVVLGNAFT 115

QY 120 AGKIIISCIPPGFAAQNISIAQATMPHVIADRVLEPIEVPLEDVRNVLFHNDNAPT- 178
Db 116 AGKIIIFAIPPNFIDNLSAAQITMCPHVIADVROLEPVNLPMPDVRNFFHNGSDSR 175

QY 179 MRLVCMLYTLRASGSSGTDPPFIAGRVLTCPSPDFSEFLVPPNVVEQTKPFVSNPLP 238
Db 176 LRLAMLYTLPLRA--NNSGDDVFTVSCRVLTRPSDFSENFVLPPTVESKTKPFTLPILT 233

QY 239 NLTLSNRVPSLIIKMMVSRDHQGVQFQNGRVTLDCQLQGTTPTSASOLCKIRGSVFHA 298
Db 234 ISEMSNRFVPVIESLHTSPTENIVVQCQNGRVTLDCGLMGTTLQLLPSQICAFRGVLTSTR 293

QY 299 NGG-----NGY---NLTELDGSPYHAFES--PAPIGFPDL--GECDMHMEASPT 340
Db 294 TSRASDAQDTPTRLFNYWHIQLDNLNGTYPDPAEDIPALGTPDFRGKV-----FGVA 348

QY 341 TQFNTGDVIKQINVKQESA---FAPHLGTIQADGLSDSVSVNTNMIAKLGWSPVSDG--H 395
Db 349 SQRPDSTTRAHEAKVDTTSGRTFKLGSLEITTESD--DFDPNQPTKP---TPVGVGVDN 404

QY 396 RGVDVPWIPRYGSLTTEAAQLAPPIYPGGEALVFEMSDPEPTAHGTNGLSVPCPTIQEF 455
Db 405 EAEFQWLSLPNTSGQFTTHMNLAPAVAPNPFGEQLLFFRSQLPSSGSRNGVLDCLVPQE 464

QY 456 FVTHFVNEQAPTRGEAALLHYLDPDTHNLGFEKLYPEGFMTCPVNSSGTGPQTLPINGV 515
Db 465 WQHFYQESAPAQTOVALRVYVNPDTGRVLFPAKHLKLGEMTIAKN--GDSPTVPPNGY 522

QY 516 FVFSWWSRFRYQLKPVGTAGPACRL 540
Db 523 RFESWNPFTYTLAPMGTGNGRRRI 547

RESULT 13
US-09-926-799-6
; Sequence 6, Application US/09926799

```

```

; GENERAL INFORMATION:
; APPLICANT: TAKEDA, NAKAZU
; APPLICANT: NATORI, KATSURO
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: KAMATA, KUNIO
; APPLICANT: SATO, TOSHINORI
; APPLICANT: SATO, SEIYA
; TITLE OF INVENTION: Detection Kit for SRSV
; FILE REFERENCE: 217039USOXPT
; CURRENT APPLICATION NUMBER: US/09/926,799
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: JP 11175928
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: JP 11-175928
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 548
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE
; US-09-926-799-6

Query Match 40.7%; Score 1178.5; DB 23; Length 548;
Best Local Similarity 45.7%; Pred. No. 6.8e-104;
Matches 258; Conservative 82; Mismatches 183; Indels 41; Gaps 14;

QY 1 MMASKDAPTNDGTSAGQLVPEANTAEPISEMPVAGAAATAAGVNMDIPWIMNNY 60
Db 1 MKMASNDAAPSNDGAAG---LVPEINN--EAMALDPVAGAAIAAPLTGQNIIDPWINNF 56

QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSLHAQMYNGWGNMKVKVLLAGNAFTA 120
Db 57 VQAPGGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFEVQVVLGNAFTA 116

QY 121 GKIIISCIPPGFAAQNISIAQATMPHVIADRVLEPIEVPLEDVRNVLFHNDNAPT-M 179
Db 117 GKIIIFAIPPNFIDNLSAAQITMCPHVIADVROLEPVNLPMPDVRNFFHNGSDSRL 176

QY 180 RLVCMLYTLRASGSSGTDPPFIAGRVLTCPSPDFSEFLVPPNVVEQTKPFVSNPLPL 239
Db 177 RLIALMYTLPLRA--NNSGDDVFTVSCRVLTRPSDFSENFVLPPTVESKTKPFTLPILT 234

QY 240 NLTLSNRVPSLIIKMMVSRDHQGVQFQNGRVTLDCQLQGTTPTSASOLCKIRGSVFHAN 299
Db 235 SEMSNRFPVPIESLHTSPTENIVVQCQNGRVTLDCGLMGTTLQLLPSQICAFRGVLTSTR 294

QY 300 GG-----NGY---NLTELDGSPYHAFES--PAPIGFPDL--GECDMHMEASPTT 341
Db 295 SRASDAQDTPTRLFNYWHVQLDNLNGTYPDPAEDIPGTLCTPDFRGKV-----FGVAS 349

QY 342 QFNTGDVIKQINVKQESA---FAPHLGTIQADGLSDSVSVNTNMIAKLGWSPVSDG--HR 396
Db 350 QRNLDSSTTRAHEAKVDTTAGRTFKLGSLEISTDSD--DFDQNPQTKF---TPVGIGVDNE 405

QY 397 GDVDPWIPRYGSLTTEAAQLAPPIYPGGEALVFEMSDPEPTAHGTNGLSVPCPTIQEF 456
Db 406 AEFQWLSLPDYSGQFTTHMNLAPAVAPNPFGEQLLFFRSQLPSSGSRNGVLDCLVPQEW 465

QY 457 VTHFVNEQAPTRGEAALLHYLDPDTHNLGFEKLYPEGFMTCPVNSSGTGPQTLPINGVF 516
Db 466 VQHFYQESAPAQTOVALRVYVNPDTGKLVFEAKHLKLGEMTIANN--GDSPTVPPNGYF 523

QY 517 FVFSWWSRFRYQLKPVGTAGPACRL 540
Db 524 RFESWNPFTYTLAPMGTGNGRRRI 547

RESULT 14
US-08-273-257-11
; Sequence 11, Application US/08273257

```

; GENERAL INFORMATION:
 ; APPLICANT: LEW, Judy F.
 ; APPLICANT: GREEN, Kim Y.
 ; APPLICANT: VALDESUSO, Jose
 ; TITLE OF INVENTION: Calicivirus capsid genes and their uses
 ; NUMBER OF SEQUENCES: 36
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend Kourie and Crew
 ; STREET: Steuart Street Tower, One Market Plaza
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: US
 ; ZIP: 94105-1493
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/273,257
 ; FILING DATE: 11-JUL-1994
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bastian, Kevin L.
 ; REGISTRATION NUMBER: 34,774
 ; REFERENCE/DOCKET NUMBER: 15280-209
 ; REFERENCE/DOCKET NUMBER: DHHS Ref. No. E-163-94/0
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 543-9600
 ; TELEFAX: (415) 543-5043
 ; INFORMATION FOR SEQ ID NO: 11:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 535 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLSCULE TYPE: protein
 ; US-08-273-257-11

Query Match 40.6%; Score 1176.5; DB 6; Length 535;
 Best Local Similarity 45.8%; Pred. No. 1e-103;
 Matches 254; Conservative 88; Mismatches 181; Indels 31; Gaps 12;

QY 1 MMASKADPTNMDGTSGAGOLVPEANTAEPISEPVAGAAATAAATAGQVNMIDPWIMNNY 60
 DB 1 MKMASNDAAAPSNDGAAG---LVPEANN-ETMALEPVAGASIAAPLTQNNVIDPWIRMF 56
 QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQYNGWGNKVKVLLAGNAFTA 120
 DB 57 VQAPNGEFTVSPRNSPGEILLNLELGPENLPFLSHLSRMNGYAGGVEQVLLAGNAFTA 116
 QY 121 GKIIISCIPPGFAAQNISIAQATMFPVHTADVRVLEPIEVPLEDVRNVLFH-NNDNAPTM 179
 DB 117 GKLVFAAIPHPPLENLSFGQITMFPVHTADVRVLEPIEVPLEDVRNVLFH-NNDNAPTM 176
 QY 180 RLVCMLYTLRASGSSGSDTPFVIAGRVLTCPSPDFSFLLVPPNVEQKTKPFSVPLNPL 239
 DB 177 RLVMALYTLPLRSNG--SGDDVFTVSCRVLTPSPDFDFNLVPPVETSKTKPTPLITI 234
 QY 240 NTLNSRVPSLTKSMVSRDHQOMQFQNGRVTLDCLOGLTTPTSASOLCKIRGSVFHAN 299
 DB 235 GELTNSRFVPPIDELTYTSPNEGVVQNGRSTLDGELLGTLQVPSNLCALRGRI-NAQ 293
 QY 300 GGNYN-----LTLDGSPYHAFES-PAPIGPDLGECOMHMEASPTTQFNGDVLIKQ-- 351
 DB 294 VPDDHHQWNLQVNTNGTFEDTEDYPAELCTPDLFLANIYGV---TSORNPNNTCRAHD 349
 QY 352 -INVQESAFAPHLGTIQADGLSDVSVNTNMTAKLGWSPVSDGHRGVDVPPWIPRYGST 410
 DB 350 GVLATNSPKFTPKLGSVILGTWEESDLNQPTRF---TPVGLFNTDHFQWALPSYSGR 406
 QY 411 LTEAQAQAPPIYPGGEAIVFMSDFPIAHGTNGLSVPCCTTPOEVTHFVNEQAPTRGE 470
 DB 407 LTLNMNLAPSVSPLPFGQLLFRSHIPLKGGTSDGADICLLPQEWIQLHYQESAPAAD 466

QY 471 AALLHYLDPDTHRLNGEFLKYPGFWMTCPVNSSGTGPOTLPINGVFVSWSRFYQLKP 530
 DB 467 VALIRTNPDTPGRVLPEAKLRQGFITVA--NSGSRPIVPPNGYFRFDSWVNOFYSLAP 524
 QY 531 VGTAGPACRLGIRR 544
 DB 525 MGTGN-----GRRR 533
 RESULT 15
 US-09-926-799-8
 ; Sequence 8, Application US/09926799
 ; GENERAL INFORMATION:
 ; APPLICANT: TAKEDA, NAOKAZU
 ; APPLICANT: NATORI, KATSURO
 ; APPLICANT: MIYAMURA, TATSUO
 ; APPLICANT: KAMATA, KUNIO
 ; APPLICANT: SATO, TOSHINORI
 ; APPLICANT: SATO, SEIYA
 ; TITLE OF INVENTION: Detection kit for SRSV
 ; FILE REFERENCE: 217039US0XPCT
 ; CURRENT APPLICATION NUMBER: US/09/926,799
 ; CURRENT FILING DATE: 2002-03-29
 ; PRIOR APPLICATION NUMBER: JP 11175928
 ; PRIOR FILING DATE: 1999-06-22
 ; PRIOR APPLICATION NUMBER: JP 11-175928
 ; PRIOR FILING DATE: 1999-06-22
 ; NUMBER OF SEQ ID NOS: 34
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 8
 ; LENGTH: 535
 ; TYPE: PRT
 ; ORGANISM: ARTIFICIAL SEQUENCE
 ; FEATURE:
 ; OTHER INFORMATION: SYNTHETIC PEPTIDE
 ; US-09-926-799-8

Query Match 40.6%; Score 1174.5; DB 23; Length 535;
 Best Local Similarity 45.0%; Pred. No. 1.6e-103;
 Matches 251; Conservative 87; Mismatches 181; Indels 39; Gaps 13;

QY 1 MMASKADPTNMDGTSGAGOLVPEANTAEPISEPVAGAAATAAATAGQVNMIDPWIMNNY 60
 DB 1 MKMASNDAAAPSNDGAAG---LVPEANN-ETMALEPVAGASIAAPLTQNNVIDPWIRNF 56
 QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQYNGWGNKVKVLLAGNAFTA 120
 DB 57 VQAPNGEFTVSPRNSPGEILLNLELGPENLPFLSHLSRMNGYAGGVEQVLLAGNAFTA 116
 QY 121 GKIIISCIPPGFAAQNISIAQATMFPVHTADVRVLEPIEVPLEDVRNVLFH-NNDNAPTM 179
 DB 117 GKLVFAAIPHPPLENLSFGQITMFPVHTADVRVLEPIEVPLEDVRNVLFH-NNDNAPTM 176
 QY 180 RLVCMLYTLRASGSSGSDTPFVIAGRVLTCPSPDFSFLLVPPNVEQKTKPFSVPLNPL 239
 DB 177 RLVMALYTLPLRSNG--SGDDVFTVSCRVLTPSPDFDFNLVPPVETSKTKPTPLITI 234
 QY 240 NTLNSRVPSLTKSMVSRDHQOMQFQNGRVTLDCLOGLTTPTSASOLCKIRGSVFHAN 299
 DB 235 GELTNSRFVPPIDELTYTSPNEGVVQNGRSTLDGELLGTLQVPSNLCALRGRI-NAQ 294
 QY 300 GGNYN-----LTLDGSPYHAFES-PAPIGPDLGECOMHMEASPTTQFNGDVLIKQ-- 346
 DB 295 VPDDHHQWNLQVNTNGTFEDTEDYPAELCTPDLFLANIYGV---TSORNPNNTCRAHD 348
 QY 347 DVTKIQVQESAFAPHLGTIQADGLSDVSVNTNMTAKLGWSPVSDGHRGVDVPPWIPR 406
 DB 349 DAVIATN---SAKFTPKLGAIQIGTWEEEDVHINQPTKPF---TPVGLFENEGRNQWTLPN 402
 QY 407 YGSLTEAQAQAPPIYPGGEAIVFMSDFPIAHGTNGLSVPCCTTPOEVTHFVNEQAP 466
 DB 403 YSGALTLMNLAPSVSPLPFGQLLFRSHIPLKGGVADVDICLLPQEWIQLHYQESAP 462

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 09:55:20 ; Search time 20 Seconds
(without alignments)
1990.084 Million cell updates/sec

Title: US-09-926-799-1
Perfect score: 2896
Sequence: 1 MMASKDAPTNDGTSGAGQ.....YOLKPVGTAGPACRLGTRRS 545

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 271938 seqs, 73030588 residues

Total number of hits satisfying chosen parameters: 271938

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

*Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1950	67.3	530	US-10-314-739-3	Sequence 3, Appli
2	272.5	9.4	669	US-10-209-507-2	Sequence 2, Appli
3	272	9.4	668	US-10-209-507-4	Sequence 4, Appli
4	109	3.8	10431	US-60-427-045-310	Sequence 310, App
5	106.5	3.7	931	PCT-US02-33645-26	Sequence 26, Appl
6	103.5	3.6	1210	US-10-258-106-1	Sequence 1, Appli
7	102.5	3.5	1210	US-10-293-017-66	Sequence 66, Appl
8	100	3.5	6879	US-60-419-463-26	Sequence 26, Appl
9	99	3.4	4961	PCT-US02-10366-64	Sequence 64, Appl
10	98	3.4	490	US-10-258-951-67	Sequence 67, Appl
11	97.5	3.4	1328	PCT-US02-39126-6	Sequence 6, Appli
12	97.5	3.4	2126	US-10-052-648A-39	Sequence 39, Appl
13	97	3.3	584	US-09-724-676-91417	Sequence 91417, A
14	97	3.3	584	US-09-724-676A-91417	Sequence 91417, A
15	97	3.3	645	US-09-724-676-91423	Sequence 91423, A
16	97	3.3	645	US-09-724-676A-91423	Sequence 91423, A
17	96.5	3.3	903	US-09-724-676-94702	Sequence 94702, A
18	96.5	3.3	903	US-09-724-676A-94702	Sequence 94702, A
19	96.5	3.3	941	US-09-724-676-94682	Sequence 94682, A
20	96.5	3.3	941	US-09-724-676A-94682	Sequence 94682, A
21	96.5	3.3	943	US-09-724-676-94698	Sequence 94698, A
22	96.5	3.3	943	US-09-724-676-94699	Sequence 94699, A
23	96.5	3.3	943	US-09-724-676-94700	Sequence 94700, A
24	96.5	3.3	943	US-09-724-676-94701	Sequence 94701, A
25	96.5	3.3	943	US-09-724-676A-94698	Sequence 94698, A
26	96.5	3.3	943	US-09-724-676A-94699	Sequence 94699, A

27	96.5	3.3	943	5	US-09-724-676A-94700	Sequence 94700, A
28	96.5	3.3	943	5	US-09-724-676A-94701	Sequence 94701, A
29	96.5	3.3	968	5	US-09-724-676-94707	Sequence 94707, A
30	96.5	3.3	968	5	US-09-724-676A-94707	Sequence 94707, A
31	96.5	3.3	981	5	US-09-724-676-94678	Sequence 94678, A
32	96.5	3.3	981	5	US-09-724-676-94679	Sequence 94679, A
33	96.5	3.3	981	5	US-09-724-676-94680	Sequence 94680, A
34	96.5	3.3	981	5	US-09-724-676-94681	Sequence 94681, A
35	96.5	3.3	981	5	US-09-724-676A-94678	Sequence 94678, A
36	96.5	3.3	981	5	US-09-724-676A-94679	Sequence 94679, A
37	96.5	3.3	981	5	US-09-724-676A-94680	Sequence 94680, A
38	96.5	3.3	981	5	US-09-724-676A-94681	Sequence 94681, A
39	96.5	3.3	1006	5	US-09-724-676-94687	Sequence 94687, A
40	96.5	3.3	1006	5	US-09-724-676A-94687	Sequence 94687, A
41	96.5	3.3	1008	5	US-09-724-676-94703	Sequence 94703, A
42	96.5	3.3	1008	5	US-09-724-676-94704	Sequence 94704, A
43	96.5	3.3	1008	5	US-09-724-676-94705	Sequence 94705, A
44	96.5	3.3	1008	5	US-09-724-676-94706	Sequence 94706, A
45	96.5	3.3	1008	5	US-09-724-676A-94703	Sequence 94703, A

ALIGNMENTS

RESULT 1
US-10-314-739-3
; Sequence 3, Application US/10314739
; GENERAL INFORMATION:
; APPLICANT: Estes, Mary K
; Jiang, Xi
; Graham, David Y
; TITLE OF INVENTION: Methods and Reagents to Detect and
; Characterize Norwalk and Related Viruses
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 801 Pennsylvania Ave., N.W.
; CITY: Washington, D.C.
; STATE: <Unknown>
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/314,739
; FILING DATE: 09-Dec-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,049
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Davis, Peter
; REGISTRATION NUMBER: 36,119
; REFERENCE/DOCKET NUMBER: 311.023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-662-0200
; TELEFAX: 202-662-4643
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3

Query Match 67.3%; Score 1950; DB 6; Length 530;
Best Local Similarity 67.0%; Pred. No. 1e-160;
Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;

Thu Jan 16 15:59:22 2003

us-09-926-799-1.rapn

QY 1 MMASKDAPTNWDGTSAGQLVPEANTAEPTSMPEVAGATAAATAGQVNMIDPWIMNNY 60
DB 1 MMASKDAPTNWDGTSAGQLVPEANTAEPTSMPEVAGATAAATAGQVNMIDPWIMNNY 60
QY 61 VOAPGEEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGVGNMKVLLAGNAFTA 120
DB 61 VOAPGEEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGVGNMKVLLAGNAFTA 120
QY 121 GKIIISCIPTPPGFAAQNISIAQATMFPHVIAVRVLEPIEVLDPVNRVLFHND-NAPT 179
DB 121 GKIIISCIPTPPGFAAQNISIAQATMFPHVIAVRVLEPIEVLDPVNRVLFHND-NAPT 179
QY 180 RLVCMLYTPLRASGSSGTPDPIAGRVLTCPSPDFSEFLVPPNVEQKTPFSPNPLPL 239
DB 180 RLVCMLYTPLRASGSSGTPDPIAGRVLTCPSPDFSEFLVPPNVEQKTPFSPNPLPL 239
QY 181 RLVCMLYTPLRGSGTG--DSFVAGRVMTCPSPDFSEFLVPPNVEQKTPFSPNPLPL 238
DB 181 RLVCMLYTPLRGSGTG--DSFVAGRVMTCPSPDFSEFLVPPNVEQKTPFSPNPLPL 238
QY 240 NTLNSRVPSPKMMVSRDHQOMVOFNGRVTLTGQLOGTTPTSASQLCKIRGSPFHAN 299
DB 240 NTLNSRVPSPKMMVSRDHQOMVOFNGRVTLTGQLOGTTPTSASQLCKIRGSPFHAN 299
QY 239 SLSNSRAPLPISSMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT----S 294
DB 239 SLSNSRAPLPISSMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT----S 294
QY 300 GGNYNLTLDGSPYHAFESPAPIGPDGLGCDWHMEASPTTQFNTGDVVIKQINVKQESA 359
DB 300 GGNYNLTLDGSPYHAFESPAPIGPDGLGCDWHMEASPTTQFNTGDVVIKQINVKQESA 359
QY 295 NGTVINLTLDGTPPFPFEGPAPIGPDGLGCDWHN---MTQFGHSSQTQYDVTDP 351
DB 295 NGTVINLTLDGTPPFPFEGPAPIGPDGLGCDWHN---MTQFGHSSQTQYDVTDP 351
QY 360 FAPHLGTIQADGLSDSVNTNMIKLGWSPVSDGHRGVDVDPWVPIRYGSTLT 419
DB 360 FAPHLGTIQADGLSDSVNTNMIKLGWSPVSDGHRGVDVDPWVPIRYGSTLT 419
QY 352 FVPHLSIQANGIG---SGNVGVLSWISPPSPHSGSQVDLWKIPNYGSSITEATHLAP 407
DB 352 FVPHLSIQANGIG---SGNVGVLSWISPPSPHSGSQVDLWKIPNYGSSITEATHLAP 407
QY 420 PIYPPGGEAIVFFMSDFPIAHCTNGLSVPCPTPOEVTHVNEQAPTRGEAALLHYLDP 479
DB 420 PIYPPGGEAIVFFMSDFPIAHCTNGLSVPCPTPOEVTHVNEQAPTRGEAALLHYLDP 479
QY 408 SVYPPGGEVLVFFMSKMP---GPGAYNLPCLLPQETIISHLASEQAPTVGEAALLHYVDP 464
DB 408 SVYPPGGEVLVFFMSKMP---GPGAYNLPCLLPQETIISHLASEQAPTVGEAALLHYVDP 464
QY 480 DTHRNLCFEKLYPEGMTCVPNSSGTPQPLTINGVFEVSVWSRYQLKPVGTAGPA-C 538
DB 480 DTHRNLCFEKLYPEGMTCVPNSSGTPQPLTINGVFEVSVWSRYQLKPVGTAGPA-C 538
QY 539 RLGIIR 544
DB 539 RLGIIR 544
QY 525 RLGLR 530
DB 525 RLGLR 530

RESULT 2

US-10-209-507-2
; Sequence 2, Application US/10209507

; GENERAL INFORMATION:

; APPLICANT: Audonnet, et al.

; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT

; FILE REFERENCE: 454313-3151.2

; CURRENT APPLICATION NUMBER: US/10/209,507

; CURRENT FILING DATE: 2002-11-12

; PRIOR APPLICATION NUMBER: 09/617,594

; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: 60/193,332

; PRIOR FILING DATE: 2000-03-30

; PRIOR APPLICATION NUMBER: France 00 01761

; PRIOR FILING DATE: 2000-02-11

; PRIOR APPLICATION NUMBER: France 99 09421

; PRIOR FILING DATE: 1999-07-16

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 2

; LENGTH: 669

; TYPE: PRT

; ORGANISM: Feline calicivirus

US-10-209-507-2

Query Match 9.4%; Score 272.5; DB 6; Length 669;

Best Local Similarity 23.4%; Pred. No. 4.1e-15;

Matches 128; Conservative 76; Mismatches 194; Indels 149; Gaps 26;

QY 13 DGTSGAGQLVPEANT-AEPISEMPVAGATAAATAGQVNMIDPWIMNNYVQAPQGEFTI; 70

DB 126 DGDSSI--TTPGCTLVGGVIAEPSAQMATATAAATGKSVDSSEW-----ESFFSFH 174
QY 71 -----SPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGVGNMKVLLAGNAFTAGKIII 125
DB 175 TSNWMTSETQGTQKILFKQSLGPLLNPFLSHLAQMYNGVGNMKVLLAGNAFTAGKIII 234
QY 126 SCIPPGF-AAQNTISIAQATMFPHVIAVRVLEPIEVLDPVNRVLFHNDNAPTMRVCM 184
DB 235 IVVPPGVDVPQVSTMLQ---YPHVLEDAQVEVIFSPDLRSTLYHLMSDSTDTTSLVM 291
QY 185 LY-----TPLRASGSSGTPDPIAGRVLTCPSPDFSEFLVPPNVEQKTPFSPNPLPL 240
DB 292 VYNDLINPYANDSNSGC---IVT--VETKPGDPFKFHLKKPEG-----S 331
QY 241 TLSNSRVPSPKMMVSRDHQOMVOFNGRVTLTGQLOGTTPTSASQL 288
DB 332 MLTHGSPSPDLIPKSSSLWTGNRYWSDITDFVIRPFVQANR-HFDEN-QETAGHSTPRF 389
QY 289 CKIRGSVFHANG---GNG-----YNTFELDGSPY--- 314
DB 390 RPTITISENGSKLGTGVATDIYVPGIPDGWPDTTTIGBELTPAGDYISITNGSGNDIATA 449
QY 315 HAFESPAPI---GFPDLGECG-----WHMEASPTTQFNT-----GDVVIKQINVKQESA 360
DB 450 NAYDSADVITNTTFRGMVYICGALQRAWGDKKISSTAFITTAKEGNTLAPSWTIDMTKI 509
QY 361 A---PHLGTIQADGLSDSVNTNMIKLGWSPVSDGHRGVDVDPWVPIRYGSTLT 412
DB 510 AVYQDTHVG-----RDVQTSDDTLALGYTGIGEGAGSNNRDSVVVRISMLPETGAR-- 560
QY 413 EAAQLAPPIYPPGGEAIVFFMSDFPIAHCTNGLSVPCPTPOEVTHVNEQAPTRGEAA 472
DB 561 -----GNHPIFYKNSIKLGYLRSIDV-----FNSOILHTSRQLS 596
QY 473 LHYLDP 479
DB 597 LHYLDP 603

RESULT 3

US-10-209-507-4
; Sequence 4, Application US/10209507

; GENERAL INFORMATION:

; APPLICANT: Audonnet, et al.

; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT

; FILE REFERENCE: 454313-3151.2

; CURRENT APPLICATION NUMBER: US/10/209,507

; CURRENT FILING DATE: 2002-11-12

; PRIOR APPLICATION NUMBER: 09/617,594

; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: 60/193,332

; PRIOR FILING DATE: 2000-03-30

; PRIOR APPLICATION NUMBER: France 00 01761

; PRIOR FILING DATE: 2000-02-11

; PRIOR APPLICATION NUMBER: France 99 09421

; PRIOR FILING DATE: 1999-07-16

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 4

; LENGTH: 668

; TYPE: PRT

; ORGANISM: Feline calicivirus

US-10-209-507-4

Query Match 9.4%; Score 272; DB 6; Length 668;

Best Local Similarity 31.0%; Pred. No. 4.5e-15;

Matches 76; Conservative 36; Mismatches 77; Indels 56; Gaps 11;

QY 23 PEANT-----APP-ISMPEVAGATAAATAGQVNMIDPWIMNNYVQAPQGEFTI----- 70

DB 133 PEOGTLVGGVIAEPSAQMATATAAATGKSVDSSEW-----EAFPSFHTSVN 177

QY 71 -SPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGVGNMKVLLAGNAFTAGKIII 129

Db 178 WSTSETQKILFKQSLGGLNPLNTHLAKLYVAMSGSIEVRESISGSGVFGKLAIVVP 237
Qy 130 PGF-AAQISIAQAATMFPHVIAADVRLVLEPIEVPLDVRNVLFHNNNDNAPTMRVLCMLV-- 186
Db 238 PGIDPVQSTMLQ---YPHVLFDAQOQVEPVITIPDLRNSLYHLMSDDTDTTSLVIMIND 294
Qy 187 --TPLRASGSSGTPDFVIAGRVLCPSDPDFSLFLVPPNVBQTKPFSVPNPLNTLSN 244
Db 295 LINPYANDSNSSGC---IVT--VERKPGDPDFKHLKPPG-----SMLTH 334
Qy 245 SRVPS 249
Db 335 GSIPS 339
RESULT 4
US-60-427-045-310
; Sequence 310, Application US/60427045
; GENERAL INFORMATION:
; APPLICANT: The Board of Trustees of the University of Arkansas
; APPLICANT: O'Brien, Timothy
; APPLICANT: Beard, John
; APPLICANT: Underwood, Lowell
; TITLE OF INVENTION: Cal25 Gene and its Use for Diagnostic and Therapeutic
; TITLE OF INVENTION: Interventions
; FILE REFERENCE: 022438.44514
; CURRENT APPLICATION NUMBER: US/60/427,045
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 314
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 310
; LENGTH: 10431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-427-045-310

Query Match 3.8%; Score 109; DB 7; Length 10431;
Best Local Similarity 18.3%; Pred. No. 34;
Matches 99; Conservative 81; Mismatches 222; Indels 138; Gaps 22;
Qy 5 SKDAPTNNMGTSAGOLVPEANTAEF-ISMPEVAGAAATAAGOVNMDPIMNNYVQA 63
Db 8743 SRTEALSGLRTSTPG---PAOSTISPELSTETITRISTPLTTGSAETITPKTCHSGAS 8799
Qy 64 PQGEFTISPNNNT---PGDILFDLQGLHNPFLSHLAQYNG-----WVGNKVK- 110
Db 8800 SOGTFTLDTSSRASPWG---THSAATHRSPHSGMTTMSRCGPEDVSNPSPRSVEKTSPP 8855
Qy 111 -VLLAGNAFTAGKII-----ISCIPPGPF- 132
Db 8856 SSLVLSAVTSPSPLYSTPSESSHSLRVTSLSFTPVMMKTTMDLDTSLPVTTSPPSMN 8915
Qy 133 --AQNISIAQATM-----FPHVIADVRVLEPI--EV 160
Db 8916 ITSDESLSATKATEMETAIAQLSENTAVTQMGTISARQEFYSYIPLGPSPKVTSPVATSS 8975
Qy 161 PLEDVRNVLFHNNNDNAPTMRVLCMLVTPLRASGSS-----SGTDPFVIAGRVLCPSDP 214
Db 8976 TIKDIVSTIIPASSIETRIEMESTSLTPTPRETSTSOEIHSAKPSVTPYKALTSATIE 9035
Qy 215 FSFLVLP-----PNVEOKTKPFSVPNPLNTLSNRVPSLKSMMVSRDHGQMVQFQNGR 270
Db 9036 DSMTQVMSRRGSPDQSTMSQDISSEVITRLTSPIKAESTEMITITQTGPGATSRGT 9095
Qy 271 VTLDQO---LOGTPTT-----SASOLCKIRG-----SVFHANGNGYNLTEL 310
Db 9096 LTLDSTFTTMSGTHSTASQGFSGHSQMTALMSRTPGDVPWLSHPSVEEASAS--FSL- 9151
Qy 311 GSPYHAFESPAPIGPDLGECDDHWEASPTTQFNTGDVTKQINV-----KQESAFAPHLG 365
Db 9152 -SPVMTSSSPVSSLTLPD-----SIHSSSLPVTSLSLTGSLVKTKTELLGTSSEPETSPPNLS 9206

Qy 366 TIOADGLSDSVNNTNMIKLGWSPVSDGHRGCDVDPWVTPRYGSTLTLEAAQLAPPIYPPG 425
Db 9207 STSAEILATTEVTTD--TEKLEMTNVVTSYTHESPSSVLA--DSVTTKATSSNGITYPTG 9263
RESULT 5
PCT-US02-33645-26
; Sequence 26, Application PC/TUS0233645
; GENERAL INFORMATION:
; APPLICANT: The Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Gao, Guangping
; APPLICANT: Roy, Soumitra
; TITLE OF INVENTION: Simian Adenovirus Nucleic Acid and Amino Acid Sequences, Vectors
; FILE REFERENCE: UPN-02677PCT
; CURRENT APPLICATION NUMBER: PCT/US02/33645
; CURRENT FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: US 60/331,951
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/366,798
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 931
; TYPE: PRT
; ORGANISM: simian adenovirus SV-1
PCT-US02-33645-26

Query Match 3.7%; Score 106.5; DB 1; Length 931;
Best Local Similarity 19.2%; Pred. No. 1.6;
Matches 105; Conservative 68; Mismatches 226; Indels 147; Gaps 24;
Qy 9 PTNMDGTSGAGOLVPEAN-----TAEPISMEPVAGAAATAAGOVNMDPIMNNYVQ 62
Db 226 PTNREG----GOATPSAQDVQNPQLQFFASTNVANTPKAVLYAEDVS-----IE 271
Qy 63 APOGEFTISPNNTPG---DILFDLQGLHNPFL-----SHLAQYNGWVGNKVKVLLA 114
Db 272 ADPTHLVKPTVTEGITSSEALLTQOAPNPNRYTAFRDNFICLMYNTSTGNKV---UA 328
Qy 115 GNAFTAGKIIISCIPPGFAAQNISIAQATMPHV-----IADVRVLEP 157
Db 329 GOASOLNAV-----DLQDRNTELSYQLMDALGDRSRYFSMMNQAVDSYDPDVRIEN 382
Qy 158 IEVPLEDRNVLFHNNNDNAPTMRVLCMLYTPLRASGSSG--TDPFVIAGRVLCPSDP 215
Db 383 HGVE-DELPNYCFPLGMA-----VTDTYSPIKVGGGNGWNEANGVFTTGERGVEIGSN- 435
Qy 216 SFLFLVPPNVEOKT-KPFSVPNPLN-----TLSNRVPSLKSMMVSRDHGQMVQFQ 267
Db 436 --MFAMENLQANLWRSFLYSNIGLYLPSDLKITPDNITLP-----ENKNTYQYM 483
Qy 268 NGRVTLDGLOQTTPTSASQLCKIRG-----FHANGNGYNLTELDSPPYHAFESPAP 323
Db 484 NGRVTPPGLVDTYVNVGARWSPDVMSINPFPNHRNAGLRYSMLLGNGRYVFEHIQVQ 543
Qy 324 GF-----PDLGECDDHWEASPTTQFNTGDVTKQINVKQESAFAPHLGTIQADGLSD 374
Db 544 KFFAIKNNLLLPSTYEWNER-----KDVNMLQSSLG---NDLRVDGASI 587
Qy 375 VSVNNTMIKLGWSPVSDGHRGCDVDPWVTPRYGSTLTLEAAQLAPPIYPPGGEATVFP 434
Db 588 RPSDINLVANF---FPM-----HNTASTLEAMLRNDTNDQSFENDYLCAAN 630
Qy 435 SDFPIAHGTNGLSVCTIPOE-----FVTHFVNEQAPTRG-----AALLHYLD 478
Db 631 MLYPIP--ANATSVISIPSRNNAAFRGSFTRLKTKETPSLGSFGDPYFYVSGSIPYLD 688
Qy 479 PDTHRN 484
Db 689 GTFYLN 694

214 ---DFSEFLFYPNVQKTKPFSVPNPLNLTLS-----NSRVPSLIKSMVVS 257
 874 LRTTSSNSLVP--VQDQHQPIIPDTPSPVPVSVITIRSDTDEEDNKYKPS---SSGLKP 929
 258 RDHGQWQFQNGRVTLDGLOQTTPTSASQCKIRGVSFVHANGN---GYNLTLDGSPY 314
 930 RSN--VISYVTVNDSPDSLSLSPYSTDTLSALRG-----NSGSVLEGGVRVADGTGT 982
 315 HAFESPAPGFDPDLGECDDHMEASPTQFNTGVDVQKQINVAQESAFAPHLCTIQADGLS- 373
 983 RTIIVP-PLK-TQLGDCTVATQASGLLSNKTVPASVSGSGGCCITPTGYRAQRGGTSA 1040
 374 ---DVSVNTNMIAKLGWVSPVSDHGRGDVDPWVIPRYGSTLTAAQAPPIYPGFGFA 429
 1041 AQLNLISQNOSSA-----APTQSRSSNPAP-----RROQAFVAP-----LSQA 1080
 430 IVFFMSDFPIAHGTNG---LSVPCTIPQEFVHFVNEQAPTRGEA-----ALLHYLDPD- 480
 1081 PYTFQHGSP-L-HSTGHPHLAPAPALPSQ--AHLTYAAPTSAALGSTSSIAHLFSPOG 1137
 481 THRNLGFEFKLYPEGFMTCPVNSSGTGPTQL 510
 1138 SSRHAAATTHPSTLTVHQVPVS--VGPSLL 1165

RESULT 7
 US-10-293-017-66
 ; Sequence 66 Application US/10293017
 ; GENERAL INFORMATION:
 ; APPLICANT: Shamloo, Mehrdad
 ; APPLICANT: Chin, Daniel J.
 ; APPLICANT: Von Schack, David
 ; APPLICANT: Urfer, Roman
 ; APPLICANT: Paffy, Laszlo
 ; APPLICANT: Wieloch, Tadeusz
 ; APPLICANT: Gonzalez-Zulueta, Mirella
 ; TITLE OF INVENTION: Regulated Genes in the Pathophysiology of Ischemic Stroke
 ; FILE REFERENCE: AGYT-007051
 ; CURRENT APPLICATION NUMBER: US/10/293,017
 ; CURRENT FILING DATE: 2002-11-12
 ; PRIOR APPLICATION NUMBER: 60/384,522
 ; PRIOR FILING DATE: 2002-05-31
 ; PRIOR APPLICATION NUMBER: 60/338,296
 ; PRIOR FILING DATE: 2001-11-09
 ; NUMBER OF SEQ ID NOS: 86
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 66
 ; LENGTH: 1210
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; OTHER INFORMATION: Incyte ID No: 2890544CD1
 US-10-293-017-66

Query Match 3.5%; Score 102.5; DB 6; Length 1210;
 Best Local Similarity 21.2%; Pred. No. 5.3;
 Matches 120; Conservative 62; Mismatches 246; Indels 139; Gaps 31;

9 PTNMDGTSAGQLVPEANTAEPISEMPVAGAAT-AAATAGOVNMIDPWIMNNYVQAQGE 67
 673 PVKMD---NAVPIVQAPAAQPLQIQ--SGVLTQGSCTPLMVATLHPQVA----- 717
 68 FTISPNPTGDLFDLQGLPHLNPFLSHLAQMYNGWGNMKVYKVLVLAGNAFTAGKIISC 127
 718 -TITPOYA---VPFTLSCAAGRPALVEQTAALQAWPGGTQ-QILLP-----SAWQOL--- 765
 128 IPPGFAAQNSTIAQATMFPHVITADVRVLEPIEVPLEDVNRVLFHNNNDNAPTMRVLCMLYT 187
 766 --PGVALHN-SVQPAAMPEAMGSGQ-----QLADWRNASHGNOYSTIMQOQSLLTN 815
 188 -----PLRAS-----RAGSSS-----GTDPFVIAGRVLTCPSPDFSLF----- 218
 816 HVTLATQAQLNVGVAHVVRQQOSSLPSKKNKQAPVSSKSSLEVLPQSVYSLVSSPLR 875
 219 -----FLVPPNVQKTKPFSVPNPLNLTLS-----NSRVPSLIKSMVVS 259

US-10-258-106-1
 ; Sequence 1, Application US/10258106
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE GENOMICS, INC.; YUE, Henry
 ; APPLICANT: GANDHI, Aneena R.; TRIBOULEY, Catherine M.
 ; APPLICANT: KEARNEY, Liam; GRIFFIN, Jennifer A.
 ; APPLICANT: NGUYEN, Dannel B.; BANDMAN, Olga
 ; APPLICANT: LU, Dyung Aina M.; LAL, Preeti G.
 ; APPLICANT: BURFORD, Neil; KHAN, Farrah A.
 ; APPLICANT: WALIA, Narinder K.; YAO, Monique G.
 ; APPLICANT: PATTERSON, Chandra; BURRILL, John D.
 ; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.
 ; APPLICANT: RECIPON, Shirley A.; LU, Yan
 ; APPLICANT: POLICKY, Jennifer L.; THORNTON, Michael B.
 ; APPLICANT: TANG, Y. Tom; HAFALIA, April J.A.
 ; APPLICANT: ELLIOTT, Vicki S.; BAUGHN, Mariah R.
 ; APPLICANT: WALSH, Roderick T.; RAMKUMAR, Jayalaxmi
 ; APPLICANT: BOROWSKY, Mark L.; AU-YOUNG, Janice K.
 ; APPLICANT: HILLMAN, Jennifer L.; GURURAJAN, Rajagopal
 ; TITLE OF INVENTION: HUMAN KINASES
 ; FILE REFERENCE: PI-0076 USN
 ; CURRENT APPLICATION NUMBER: US/10/258,106
 ; CURRENT FILING DATE: 2002-10-18
 ; PRIOR APPLICATION NUMBER: PCT/US01/12992
 ; PRIOR FILING DATE: 2001-04-20
 ; PRIOR APPLICATION NUMBER: US 60/199,021
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/200,226
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/202,339
 ; PRIOR FILING DATE: 2000-05-05
 ; PRIOR APPLICATION NUMBER: US 60/203,505
 ; PRIOR FILING DATE: 2000-05-11
 ; PRIOR APPLICATION NUMBER: US 60/205,654
 ; PRIOR FILING DATE: 2000-05-18
 ; PRIOR APPLICATION NUMBER: US 60/207,739
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 60/208,795
 ; PRIOR FILING DATE: 2000-06-01
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 1
 ; LENGTH: 1210
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No: 2890544CD1
 US-10-258-106-1

Query Match 3.6%; Score 103.5; DB 6; Length 1210;
 Best Local Similarity 20.9%; Pred. No. 4.4;
 Matches 119; Conservative 65; Mismatches 241; Indels 145; Gaps 33;

9 PTNMDGTSAGQLVPEANTAEPISEMPVAGAAT-AAATAGOVNMIDPWIMNNYVQAQGE 67
 673 PVKMD---NAVPIVQAPAAQPLQIQ--SGVLTQGSCTPLMVATLHPQVA----- 717
 68 FTISPNPTGDLFDLQGLPHLNPFLSHLAQMYNGWGNMKVYKVLVLAGNAFTAGKIISC 127
 718 -TITPOYA---VPFTLSCAAGRPALVEQTAALQAWPGGTQ-QILLPS---TWOQL----- 765
 128 IPPGFAAQNSTIAQATMFPHVITADVRVLEPIEVPLEDVNRVLFHNNNDNAPTMRVLCMLYT 187
 766 --PGVALHN-SVQPAAMPEAMGSGQ-----QLADWRNASHGNOYSTIMQOQSLLTN 815
 188 -----PLRAS-----GSSSGTDPFVIAGRV--LTCPS 213
 816 HVTLATQAQLNVGVAHVVRQQOSSLPSKKNKQAPVSSKSSLD--VLPSQVYSLVSSPLR 873


```
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-10366-64

Query Match      3.4%; Score 99; DB 1; Length 4961;
Best Local Similarity 19.3%; Pred. No. 85;
Matches 129; Conservative 80; Mismatches 254; Indels 204; Gaps 29;

QY 12 MDGTSG----AGQVPEAN-----TAEPISMEPVAGATAAATAGOVANNIDPWNNY 60
Db 2911 LDNDGTGWSVASSLSDLNQNFITVTAKDGNPPLSQATVHTVTEENVHHPFQSOSH 2970
QY 61 VOAPQGEFTISPNTPGDIL-----FDLQGLPFLNPFLLSHLAQMYNGW 103
Db 2971 MSA-----TIPESHIGSIVRTVSARDRAAMGLIKYSISSGNEEGIF-----AINSS 3019
QY 104 VGNMKYKVLLAGNAFTAGKIIISCIPGFAAQNISIAQATMFPHVIADRVLEPIE 163
Db 3020 TGIITLAKALDYELCKKHEMTISAIDGGWART---GYCSVTNNVI-DVNDNSPVFLSD 3075
QY 164 DVNVLPHNNDNAPTMRLVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPP 223
Db 3076 YFPTVL-----ENAPSGTIV---IHLNATDADSGTNAVI-----AYTVOSSD-SDLFVIDP 3122
QY 224 NV-----EQK-----TKPFSVP-----NLPPLNTLSNRVPSLIKSMV 256
Db 3123 NTGVIITQGLDFETKOSYHLTVKAFNVDPDERCSFATVNIQLKG-TNEYVPRFVSKLYY 3181
QY 257 -----SRDHGQWQVQNGRVTLDGQL--- 277
Db 3182 FEISEAPKGTIVGEVAFSDRLDGTGEVHYLIFGNSRKKGFQINKKQIIVSGILD 3241
QY 278 ---QGTTPTSASQCKIRG-----SVFHANGG-----NGYNLTELDGP 313
Db 3242 KEERVSUKVLAKNFGSIRGADIDEVTNNVTLDANDPPIFTLNISVQISEGVPIGTHVT 3301
QY 314 -YAFESPAPIGPFDGECQWDM-EASPTQFNTGDVVIQINVKQE- 360
Db 3302 FVSAFSDSDS---IPSWRSRFSYFTGSGNENGAFSINPQTGOITVTAELDRETLPYNLSVL 3358
QY 361 APHLGTTQADGLSDVSNTNNIAKLGWSPVSDGHRGDVDPWPIRGYSTLTLEAAQLAPP 420
Db 3359 AVDSGTSPAGSASLLVLTLEDINDNGPMLTVSEGEVME-----NKRPGTLVMTLQSTDP 3412
QY 421 IYPFGGEAIVFFMSDRP-----IAHGTING-LSVPCITIPQEFVTHFV-----NEQAPTRG 469
Db 3413 DLPNQGPFTYLLSTGPATSYFSLTAGVLTSTREIDREQIADFLSVVTKDSGVPMQS 3472
QY 470 EAALLHYLDPDTHNLGCEFKLYPEGFTCVPNSSSGTGQTLPINGVFEVFSWYSRFRYQLK 529
Db 3473 STGTVHITVIDQDN-----PSQSR-----VEIFVNYGYNLFP 3508
QY 530 PVGTAGP 536
Db 3509 ILGSVKP 3515

RESULT 10
US-10-258-951-67
; Sequence 67, Application US/10258951
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdock, Paul R.
; APPLICANT: Rizvi, Safia K.
; APPLICANT: Smith, Randall F.
; APPLICANT: Xiang, Zhaoqing
; APPLICANT: Kabnick, Karen
; APPLICANT: Lai, Ying-Ta
; APPLICANT: Xie, Qing
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50025
; CURRENT APPLICATION NUMBER: US/10/258,951
; CURRENT FILING DATE: 2002-10-28

; PRIOR APPLICATION NUMBER: PCT/US01/13360
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/199,963
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/203,336
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: 60/207,087
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/207,546
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 67
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-258-951-67

Query Match      3.4%; Score 98; DB 6; Length 490;
Best Local Similarity 22.8%; Pred. No. 3.5;
Matches 104; Conservative 56; Mismatches 162; Indels 134; Gaps 27;

QY 101 NGWGNMKYKVLLAGNAFTAGKIIISCIP-----PGFAAQNISIAQATMFPHVIADRV 155
Db 51 NGWTASDPVH---GYWFRAGDHVSRNIPVATNNPARAVQEEETDRRF-----HLLGDPQ-N 101
QY 156 EPIEVPLEDVR-----NVLFHNNDNAPTMRLVCMLYTPLRASGSSGTDPPF 201
Db 102 KCTLSLSTRTRESADAGTYVFCVERGNMKWNYKYDQLSVNVTALTTHMP-----TF 150
QY 202 VIAG-----RVLTCPSPDFSLFLVPPNVEQTKP-----FVSPNPLPLNTLSNRVP 248
Db 151 SIPGTLESQHPRNLTCG-----VPWACEQGTPTITWMGASVSSLD-----P 192
QY 249 SLIKSMVMS-----RDHGQWVQNGRVTLDGQLGQTTPTSASQL-----CKIRGSVFH 297
Db 193 TITRSSMLSLIPQPDHGTSLTCQ---VTLPG--AGVTWTRAVRLNLSYPPQNLTTWTFQ 247
QY 298 ANG-----GNGYNLTELDGPYH---AFESPAPIGPFDGECQW---HMEASPTQFNT 345
Db 248 GDGTASTTLTRNGSALSLEQSLHLVCAVDSNPP-----ARLSWTGSLTLPSPSSNNL 301
QY 346 GDV-IKQINVKQESAFAPH-----LGTQADGLSDVSNTNNIAKLGWSPVSDGHRGDVD 400
Db 302 GYLELPRVHVKDEGEFTCRANPLGS-QHLSLS-LSLQNEYTKMKRPISGVTILGATGGAG 359
QY 401 P-----WIPRYGSTLTLEAAQLAPPYPPGF--GEAIVFFMSDFPIAHGNTNLS 447
Db 360 ATALVFLYFCIIFVVVR--SCRKKSARPAVGVGDTGMEDANAVRGSASQGLPTESWKD-G 416
QY 448 VPCITIPQEFVTHFVNEQAPTRGEAALLHYLDPDTHR 483
Db 417 NPLKKPPPAV-----APSSGEGEGELHYATLSFHK 445

RESULT 11
PCT-US02-39126-6
; Sequence 6, Application PC/TUS0239126
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: KABLE, Amy E.
; APPLICANT: CHIEN, David
; APPLICANT: WILSON, Amy D.
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: GORVAD, Ann E.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: JIN, Pei
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: MARQUIS, Joseph P.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: CHAWLA, Narinder K.
```

```

; APPLICANT: LEHR-MASON, Patricia M.
; APPLICANT: KHARE, Reena
; APPLICANT: LEE, Sally
; APPLICANT: HAWKINS, Phillip R.
; APPLICANT: BECHA, Shanya D.
; APPLICANT: LEE, Soo Yeun
; APPLICANT: SPRAGUE, William W.
; APPLICANT: ZEBARJADIAN, Yeganeh
; TITLE OF INVENTION: KINASES AND PHOSPHATASES
; FILE REFERENCE: PF-1315 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/39126
; PRIOR APPLICATION NUMBER: US 60/340,235
; PRIOR FILING DATE: 2002-12-06
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/343,007
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/343,546
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US 60/354,388
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/357,675
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 1328
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7503214CD1
; PCT-US02-39126-6

Query Match      3.4%; Score 97.5; DB 1; Length 1328;
Best Local Similarity 20.0%; Pred. No. 17;
Matches 108; Conservative 57; Mismatches 207; Indels 169; Gaps 26;

QY   84  QLGPHLNPFLESH---LAQYNGW-VGNMKVKVLLA---GNAFTAGKIISCIPPGF----132
Db   472  KIGVRLNLLGRKSLSEKMNNDVQGFSSVLAHDVGVQAQAERFLKPKPPVWYLR531

QY   133  AQAQNTSIA---QATMPHPHVIADRVLEPTEPLEVDVNNVLFHNDNAPMRL-VCMLYTP 188
Db   -532  LVONLLIRRFKKTIEHSPPROELNFWDDIIFEATNEVT--NGLRFPVLVIEPTKVYQP 589

QY   189  LRASGSSCTDPEVIAGRVLTCPSPDFSEFLVPPNVEOKTKPFSVPNPLNLTLSNRYP 248
Db   590  SYVSINNEAEERTV-----SLWHSPT---EMQMHEWNTASSIKGIRFF 632

QY   249  SLIKSMVSRDHGMVQFONGRVTLDGQLQGTTPTSASOLCKIRGSVFFH-ANG-----GN 302
Db   633  SLVKEMITWT-----AGSTVELEGETDGT-----LEYEYDHANGERVVLC 675

QY   303  G-----YNLTELD---GSPYH-----315
Db   676  GTYGIYVYAGRDLNSQVRIAIAKEIPERDSRYSQPLHDEIALHKYLRNITVQYLGVSSENG 735

QY   316  ---AFESPAPIG-FPDLGECDDHMEASPTQFTNGDVIKQINVKQESAFAPHLGTIQADG 371
Db   736  YIKIFMEQVPGSLSALLSKSGKPKPEPIKTYTKOILEGLKYLHENQIVHR--DIKGD- 792

QY   372  LSDVSVNT-----NMIAKLGWVSP-----VSDGHRGDVDVPWIP 405
Db   793  --NLVNTYSGVVKISDFGTSKRLAGVNPCTETFTGTLQYMAPEIIDQGRGYPGAPADIW 850

QY   406  RYGSTLLEAAQLAPPIYPFCGEAIVFMSDPPIAHGTNGLSVPCPTIPOEFVTHFVNEQA 465
Db   851  SLGCTIIEMATSKPPFHELGEPQAAFMFKYGMFKI-H-----PE-----IPEAL 892

QY   466  PTRGEAALLHYLDPDTHRNLFGEFKLYPEGFMTCVPPNSS-----GTCPQ-----TLPIN 513
Db   893  SAEARAFILSCFEPDPHKRATTAEELLREGFLRQVNVKGNKKNRIAFKPSGPGVWVLPALPTQ 952

; APPLICANT: Anderson, David
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Steven
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gunther, Erik
; APPLICANT: Kekuda, Ramesh
; APPLICANT: MacDougall, John R.
; APPLICANT: Mehraban, Foad
; APPLICANT: Patturajan, Meera
; APPLICANT: Rothenberg, Mark
; APPLICANT: Shinkets, Richard
; APPLICANT: Smithson, Glennda
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Stone, David J.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION: USING THE SAME
; FILE REFERENCE: 21402-250 (CURA-550)
; CURRENT APPLICATION NUMBER: US/10/052,648A
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,920
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/284,549
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/303,229
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,605
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/269,098
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/264,159
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/271,855
; PRIOR FILING DATE: 2001-02-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 2126
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-10-052-648A-39

Query Match      3.4%; Score 97.5; DB 6; Length 2126;
Best Local Similarity 20.5%; Pred. No. 33;
Matches 110; Conservative 57; Mismatches 161; Indels 209; Gaps 29;

QY   15  TSCAGOLVPEANTAEPISEMPVAGAATAAGQVNNIDPWIMNVTYVQA-----63
Db   1132  TTSTGVVSSGLTTLVSETPTLSSAVSSSTAPAVTVS--TTSQPQVAFSTSGSIASSTG 1189

QY   64  --PQGEFTISPNNTPGDILFDLQGLHLPNPFSLHQAQMYNGWGNMKVKVLLAGNAFTAG 121
Db   1190  SPPSGTFTTTGTTVSSVAV-----PNAKPPTVLLQQ-----VAGN--TAG 1228
```

Db 436 EMPASOSKACETQPKTQSVASANQVSRSTESPOGLPTVQRENAAKVLNPPAPAGD----- 489

QY 332 DHMEASPTQFNTGDIKQINVKQESAFAPHLGTIQADGLS 373
 Db 490 --HARLDGLSKMDTERKAASANHVKPK-----PELTSTIQGPAS 525

RESULT 14
 US-09-724-676A-91417
 ; Sequence 91417, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724, 676A
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 91417
 ; LENGTH: 584
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-724-676A-91417

Query Match 3.3%; Score 97; DB 5; Length 584;
 Best Local Similarity 20.9%; Pred. No. 5.5;
 Matches 84; Conservative 55; Mismatches 163; Indels 100; Gaps 18;

QY 15 TSGAGQLVPEANTAEPTSMPEVAGAAATAAGQVNMIDPWIMNNYVQAPQGEFTISPN 74
 Db 181 SSMAGKAVAITQSPSSVRLPPAAAEHSPQTAAGMPVASP---HPDQEQKQITLQP-- 235
 QY 75 TPGDILFDLQGLPHLNPFLSHL---AQMYNGWVGNM---KVKVLLAGNAFT-----AG 121
 Db 236 TPG-----LSPHPLHLF-SHLPLHSQQQSRTPYNNVMPVGGIHVVPAAGLTYSFVPLQAG 288
 QY 122 KIIISCIPP-----GFAAQNISIAQATMPPHVIADRVLEPIEVPLEDVRNVLFHNN 173
 Db 289 PVQLT-IPAVSVVHRTLTGTHRTVTVEGTTNPAGVAELSSVWPC-IPIGQIRVPGQLN- 345
 QY 174 DNAPTRLVCMLYTPLRASGSSGDPFVIAGRVLTCPSPDFSLFLVPP----- 223
 Db 346 -----LSTPGLQSLPSLSMETVNVGLANTNNAPO-----VHPPGLALNAVGLQ 389
 QY 224 ----NVEQTKPFSPNPLNTLSNSRPSLIKSMVSRDHGQMVQFONGRVTLTGQLOG 279
 Db 390 VLTANPSSOSSPAQAHIPGLQILNIALPTLIPSV-----SQVAVDAQAP 435
 QY 280 TTPTSASQLCKIRGVSFHANGNGYNLTTEL-DGSPYHAFES-----PAPIGFDDLGE 331
 Db 436 EMPASOSKACETQPKTQSVASANQVSRSTESPOGLPTVQRENAAKVLNPPAPAGD----- 489
 QY 332 DHMEASPTQFNTGDIKQINVKQESAFAPHLGTIQADGLS 373
 Db 490 --HARLDGLSKMDTERKAASANHVKPK-----PELTSTIQGPAS 525

RESULT 15
 US-09-724-676-91423
 ; Sequence 91423, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724, 676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 91423
 ; LENGTH: 645
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-724-676-91423

QY 122 KIIISCIP-----PCFAAQ-NISIAQATMEPHVIADRVLEPIEVPLEDVRNVLFHNN 175
 Db 1229 VAIVTSVSTTFFPAMASOPSLUGSSSAP-TLAETVV-----VSAHSIDK 1274
 QY 176 A---PTMLRVLCMLYPLRASGSSGT-----DPFVIAGRVLTCPSPDFSLFL 220
 Db 1275 ASHSSTAGLGLSFCAP--SSSSSGTAVSSVSQPGIVHPLVISAIASTP-----VL 1325
 QY 221 VPPNVEQTKPF-SVPNPLNTLSNSRPSLIKSMVSRDHGQMVQFONGRVTLTGQLOG 279
 Db 1326 POPAVPTSTPLLPQVNPVPLPVANPAVQOTLIHSOPQAPALL----- 1370
 QY 280 TTPTSASQLCKIRGVSFHANGNGYNLTTELDSQSYHAFESPAPIGFDDLGECDHMEASP 339
 Db 1371 --PNQPHTHC-----PEMDADT---QSKAP-GIDD----- 1394
 QY 340 TTFQNTGDIKQINVKQESAFAPH--LGTIQADGLSDSVSVNTMIATKLGWVSPVSDGHRG 397
 Db 1395 -----IKTLEBKRLSLFSEHSSSGTQHA-----SVSLETPLVYET--VTP-----G 1433
 QY 398 DVDPWIPRYGSLTEAQAOLAPPIYPGGEAIVFPMDFPIAHGTNGLSV-PCTIPOEF 456
 Db 1434 IPTTAVAPSKLMTSTSTCLPP-----TNLPL--GTAGMPVMPGTPTGQV 1476
 QY 457 VTHFVNEQAPTRGEAALLHYLDPDTHRNLTGFEKLYPEGEWTCVPNSSGTG--PQTL 511
 Db 1477 ST-----PGTHAS-----APASTATGAKPGTTP 1499

RESULT 13
 US-09-724-676-91417
 ; Sequence 91417, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724, 676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 91417
 ; LENGTH: 584
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-724-676-91417

Query Match 3.3%; Score 97; DB 5; Length 584;
 Best Local Similarity 20.9%; Pred. No. 5.5;
 Matches 84; Conservative 55; Mismatches 163; Indels 100; Gaps 18;

QY 15 TSGAGQLVPEANTAEPTSMPEVAGAAATAAGQVNMIDPWIMNNYVQAPQGEFTISPN 74
 Db 181 SSMAGKAVAITQSPSSVRLPPAAAEHSPQTAAGMPVASP---HPDQEQKQITLQP-- 235
 QY 75 TPGDILFDLQGLPHLNPFLSHL---AQMYNGWVGNM---KVKVLLAGNAFT-----AG 121
 Db 236 TPG-----LSPHPLHLF-SHLPLHSQQQSRTPYNNVMPVGGIHVVPAAGLTYSFVPLQAG 288
 QY 122 KIIISCIPP-----GFAAQNISIAQATMPPHVIADRVLEPIEVPLEDVRNVLFHNN 173
 Db 289 PVQLT-IPAVSVVHRTLTGTHRTVTVEGTTNPAGVAELSSVWPC-IPIGQIRVPGQLN- 345
 QY 174 DNAPTRLVCMLYTPLRASGSSGDPFVIAGRVLTCPSPDFSLFLVPP----- 223
 Db 346 -----LSTPGLQSLPSLSMETVNVGLANTNNAPO-----VHPPGLALNAVGLQ 389
 QY 224 ----NVEQTKPFSPNPLNTLSNSRPSLIKSMVSRDHGQMVQFONGRVTLTGQLOG 279
 Db 390 VLTANPSSOSSPAQAHIPGLQILNIALPTLIPSV-----SQVAVDAQAP 435
 QY 280 TTPTSASQLCKIRGVSFHANGNGYNLTTEL-DGSPYHAFES-----PAPIGFDDLGE 331

```
Query Match      3.3%; Score 97; DB 5; Length 645;
Best Local Similarity 20.9%; Pred. NO. 6.4;
Matches 84; Conservative 55; Mismatches 163; Indels 100; Gaps 18;

QY 15 TSGAGOLVPEANTAEPISEVPVAGATAATAAGQVNMIDPWIMNNYVQAPOQEFTISPN 74
Db 242 SMAGKAVAITSPSSVRLPFAAAHSPQTAAGMPSVASP---HPDPQEKQOITLQP-- 296

QY 75 TPGDILFDLQLGPHLNPFLSHL---AQMYNGWVGNN---KVKVLIAGNAFT-----AG 121
Db 297 TPG-----LSPPHTHLF-SHLPLHSQQQSRTPYNNMPVGGIHVVVPAGLTYSTFVPLQAG 349

QY 122 KIIISCIPP-----GFAAONISIAQATMPHPHVIADVRLPIEVPLEDYRNVLPHNN 173
Db 350 PVQLT-IPAVSVVHRTLGLGTHRTVTVEVSGTTNPAGVAELSSVVC-IPIGQIRVPGLQN- 406

QY 174 DNAPTMRLVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPP----- 223
Db 407 -----LSTPGLQLSLPSLMSMETVNIIVGLANTNMAPO-----VHPPGLALNAVGLQ 450

QY 224 ----NVEQKTKPFSPVNPPLNTLSNRVPSLIKSMVSRDHGMVQFQNGRVTLDDQLQG 279
Db 451 VLTANPSSQSSPAPQAHIPGLQILNIALFTLIPSV-----SQVAVDAQGAP 496

QY 280 TTPTASQLCKTRGVSFHANGNGYNLTTEL-DGSPYHAFES-----PAPIGFPDLGEC 331
Db 497 EMPASQSKACETQPKQTSVASANQVSRTESPQGLPTVQRENAKVLNPPAPAGD----- 550

QY 332 DHMEASPTTQNTGDKVIQINVKQESAFAPHLGTIQADGLS 373
Db 551 --HARLDGLSKMDTEKAASANHVKPK-----PELTSIQGPAS 586
```

Search completed: January 16, 2003, 10:00:02
Job time : 37 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 09:53:05 ; Search time 21 Seconds
(without alignments)
2494.917 Million cell updates/sec

Title: US-09-926-799-1
Perfect score: 2896
Sequence: 1 MMASKDAPTNMDGTSGAGQ.....YOLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1974.5	68.2	546	2 B37491	major capsid prote
2	1946	67.2	530	2 B37471	capsid protein - N
3	1223	42.2	542	2 S60616	capsid protein - h
4	1151.5	39.8	539	2 S40111	capsid protein - h
5	316	10.9	2344	1 RWRWH	genome polyprotein
6	315	10.9	576	2 A3982	capsid protein - E
7	311.5	10.8	2344	2 S5399	genome polyprotein
8	307.5	10.6	2344	2 S64740	genome polyprotein
9	300.5	10.4	702	1 A48562	coat protein - San
10	277.5	9.6	668	2 JQ2354	capsid protein - f
11	276	9.5	671	1 VCMWF9	coat protein - fel
12	274	9.5	668	1 VCMWF6	coat protein - fel
13	269	9.3	668	1 VCMWF5	coat protein - fel
14	264	9.1	703	1 C48562	coat protein - San
15	261.5	9.0	668	2 JQ2356	capsid protein - f
16	147	5.1	2206	2 S3822	genome polyprotein
17	146.5	5.1	2205	1 GNNY2W	genome polyprotein
18	145.5	5.0	2207	1 GNNY5P	genome polyprotein
19	145	5.0	2206	1 GNNY4P	genome polyprotein
20	143.5	5.0	2194	1 GNNY7	genome polyprotein
21	141.5	4.9	2207	2 S09553	genome polyprotein
22	139.5	4.8	2206	1 GNNY27	genome polyprotein
23	137.5	4.7	2209	1 GNNY3P	genome polyprotein
24	136.5	4.7	2209	1 GNNY2P	genome polyprotein
25	135.5	4.7	2207	1 GNNY1P	genome polyprotein
26	133.5	4.6	2179	1 GNNY4	genome polyprotein
27	128.5	4.4	613	2 T35828	acetolactate synth
28	125	4.3	3473	1 A46112	genome polyprotein
29	125	4.3	3473	2 S27927	polyprotein - rice

RESULT 1

B37491
major capsid protein [similarity] - Southampton virus
N:Alternate names: orf2 protein
C:Species: Southampton virus
C:Date: 22-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 18-Aug-2000
C:Accession: B37491
R:Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
Science 259, 516-519, 1993
A:Title: Sequence and genome organization of a human small round-structured (Norwalk-
A:Reference number: A37491; MUID:93142023; PMID:8380940
A:Accession: B37491
A:Status: not compared with conceptual translation
A:Molecule type: nucleic acid
A:Cross-residues: 1-546 <LAW>
A:Cross-references: GB:L07418; MID:g1236787; PIDN:AAA92984.1; PID:g295114
A:Note: sequence extracted from NCBI backbone (NCBIP:123458)
C:Superfamily: human calicivirus capsid protein
C:Keywords: glycoprotein
F:303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

ALIGNMENTS

Query Match	68.2%	Score	1974.5	DB 2	Length	546			
Best Local Similarity	66.6%	Pred. No.	6.7e-139						
Matches	367	Conservative	75	Mismatches	96	Indels	13	Gaps	7
QY	1	MMASKDAPTNMDGTSAGOLVPEANTAPISMEPVAGATAAATAGQVNMIDPWIMNNY	60						
Db	1	MMASKDAPQSGADGAGAGOLVPEVNTADPLMEPVAGPTTAVATAGQVNMIDPWVNNF	60						
QY	61	VOAPOGEFTISPNNTPGDITFDLQGLPHLNFSLHQAQYNGVGNMVKVLLAGNAFTA	120						
Db	61	VOSPOGEFTISPNNTPGDITFDLQGLPHLNFSLHLSQYNGVGNMVRVILLAGNAFSA	120						
QY	121	GKIIISCIPPGFAAQNISIAQATMPHVIADVRVLEPIEVLDEVRNVLPHNNDNAPTMR	180						
Db	121	GKIIIVCCVPPGFTSSSLTAAQATLFPHVIADVRTLEPIEMPLEDRNVLHYHTNDQPTMR	180						
QY	181	LVCMLYTPURASGSSGTDPFVIAGRVLCPCSPDFSLFLVPNNVQKTKPFSVPNPLN	240						
Db	181	LVCMLYTPURTGSGGNSDSFVAVGRVLTAPSSDFSFLFLVPPTIEQKTRAFVTPNPLQ	240						
QY	241	TLSNSRVPSLIISMVSRDHQVQFQNGRVTLDCGLQGTTPTSASQLCKIRGVSVPFANG	300						
Db	241	TLSNSRVPSLIQGMILSPASQVQVQFQNGRCRLDGLLGTTPATSGQLFVRGKI--NOG	298						
QY	301	GNLYNLTELDGSPYHAFESPAPIDGRFDLCECDHMEASPT-TQFNTPGDVTKQINVKOE-S	358						
Db	299	ARTNLNTEVDGKPPFMAFDSPAPVGPFDGKCDHMRISKTPTNNTGSGDPMRSVSVQTNVQ	358						
QY	359	AFAPHLGTTIQADGLSDSVSYNTNNIAKLGWVSPVSDGHRGDVDPWIPRYGSTLTAAQJIA	418						

A:Molecule type: genomic RNA
A:Residues: 1-2344 <MEY>
A:CROSS-references: GB:M67473
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: nucleotidylyltransferase

Query Match 10.9% Score 316; DB 1; Length 2344;
Best Local Similarity 28.3%; Pred. No. 1.le-14;
Matches 132; Conservative 60; Mismatches 172; Indels 102; Gaps 21;

Qy 12 MDGTSAGQLVPEANTAEPPIS-----MEEPVAGNATAA-----ATAG----- 47
Db 1766 MEKARAAAPGEAACTATTASVPGTDDMDPGVVAVTTSVITAENSSASIATIGCGPQP 1825

Qy 48 QVNMDIPIMNNYYQAPOGETFISPNTTPGDILFDLQLGPLHNPFLSHLAQMNGWGCNM 107
Db 1826 QVDQETWRTFNY----NDVTWSADAPGSLLTVQUISPONNFPTAVLSOMYAGWAGGM 1882

Qy 108 KVKVLLAGNAFTAGKIISCIPPFAAQNISIA---QATMPHPHYADVRLIEPIEVLPLED 164
Db 1883 QRFRIVAGSGVEGGRLRAVIPGP-----IEIOPGLEVRQPFPHVIDARSLEPVTITPD 1937

Qy 165 VRNVLFH--NDONAPTMRUVCMLYTPL--RASGSSSGLTDFFVIAGRVLTCPSDFSLFLV 221
Db 1938 LRPNMYHTDPGLGVLTTLVSVMNLINPFEGST\$-----AIQTVTETRSEDDEFVMIR 1992

Qy 222 PPNVEQRKTSPSNLPNPLNTLSNRVSLIKSMVSRDHGOWVOFQ-----NGRVT 272
Db 1993 APS---SKTVDSISFAGLLTT-----PYLVGCVGNDRNWGOVLQRPVGFGSTCNRIWN 2044

Qy 273 LDGOLQGTTTTPTSASOLCKIRGVSFHANGCYNLTEL----DGSPYHAfESP---APIG 324
Db 2045 LNGSYGHSSPRFGDIHRRSASY-SGSNATYNLOFWANAGS---AIDNPISOVAPDG 2100

Qy 325 FPDLEGCDWHMEASPQTGFNTGDVIOKVINKQSFAAPHLCGTIQADGLSDSVNTNMIAK 384
Db 2101 PPDMSFVPFNPGPIPAAGWCFCGAIWNSN-----SGAPNVTVTVA-----YE 2142

Qy 385 LGWVSPVSDGRHDVDPWVIRYSGLTEAAQ-LAPPITYPPGFEA 429
Db 2143 LGFAT---GAPGNLQP-----TTNSTGAOTVAKSIYAVVTGTA 2177

RESULT 6
A53982
capsid protein - European brown hare syndrome virus
C:Species: European brown hare syndrome virus
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 28-Jul-2000
C:Accession: A53982
J.Wirblisch, C.: Meyers, G.; Ohlinger, V.F.; Capucci, L.; Haas, B.; Thiel,
R. Virol. 68, 5164-5173, 1994
A>Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease virus
A:Reference number: A53982; PMID:94309183; PMID:7518531
A:Accession: A53982
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-576 <WIR>
A:CROSS-references: GB:U09199
C:Superfamily: human calicivirus capsid protein

Query Match 10.9% Score 315; DB 2; Length 576;
Best Local Similarity 26.2%; Pred. No. 1.7e-15;
Matches 126; Conservative 69; Mismatches 200; Indels 86; Gaps 20;

Qy 7 DAP-----TNMGDTSGAQLYPEANTAEPISMEEPVAGAATAAGOVNMTIDP---WTMN 58
Db 8 DAPGTATTASVPGTTTGCMDBPGVVASTDVVTDADNVAASVATAGICGGPQQOASPOESRWVN 67

Qy 59 NYVQAPOGETFISPNTTPGDILLDLQLGPLHNFPFLSHLAQMNGWGCMMKKVYLACNAF 118
Db 68 FFY---NDVFVTSVTDAPGSTLYSQHSQPQNPFQTVLSQMYAGWAGMQPREFIVAGSGI 124

Qy 119 TAGKIITSCEDECAAO:NISTAOTGWEDPHVIANRVLDREVDIFDNBNVIFE-NNNNN 176

Db 125 FGRRLVCAIIPGIIQPCLEVRQ----PPHVVIDARSLPEVTTITMPDLRPEMYHPTGDGP 181
QY 177 PFMRLVCMLYPLRASGSSGTDPE----VIAGRVLTCPSPDFSEFLVPPNVEQTKP 231
Db 182 LVPTLWVSNNL-----INPEGGTTSAIQVTVETRPSEDFEFLIRAPS--SKTVD 231
QY 232 FSVPNLPLNTLSNRVPSLIKSMVSRDHGMVQFQ-----NGRVTLDGQLQCTTP 282
Db 232 SVNPSWLLTT-----PVLTGAGSDNRWCAPIVGLQPVPGGFTSNRRHWNMGSTYGWSS 285
QY 283 TSASOLCKIRGSVFHANGGNGVNLDELGSPYHAFESP-----APIGFDLGECDWHMEAS 338
Db 286 PRFDDIDHSGNVSYPTSATNTIETWYANAGTATTNPSINIAIPDGFDMGAIPF-----S 341
QY 339 PITQFNTGDIKQINVKQESAFAPHLGTIOADGLSDVSVNTNMIKLGWVSPVSDGRGD 398
Db 342 GTT-IPTGAWVFGQVWNASNGTPIVIGVQA-----YELGF-----ANGAPSS 383
QY 399 VDPWVIPRYGSLTEAAQL-APPYPPGFE-----AIVFFMSDFPIANGTNGLSVPCPTIP 453
Db 384 IRP-----VTTTGAQLVAKSIYGAIVQAQNSAGIIFLSKGMVS--TPGVAATYTP 434
QY 454 Q 454
Db 435 Q 435
RESULT 7
S55399
genome, polyprotein - rabbit hemorrhagic disease virus (isolate B589)
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate B589
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999
C:Accession: S55399
R:Rossi, C.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55399
A:Accession: S55399
A:Molecule type: genomic RNA
A:Residues: 1-2344 <ROS>
A:Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641
A:Experimental source: isolate B589
A:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: polyprotein
Query Match 10.8%; Score 311.5; DB 2; Length 2344;
Best Local Similarity 27.4%; Pred. No. 2.3e-14;
Matches 127; Conservative 60; Mismatches 172; Indels 105; Gaps 20;
QY 17 GAGQLVPEANTAEPISMPEVAGAAT-----AAATAG-----QV 49
Db 1768 GKARTAPOGEAAGATTATASVPGTTTDLGDPGVATTSVVTAENSSASIAIATAGIGPPQV 1827
QY 50 NNIDPWIMNNVQAPQGEFTISPNTPGDILFDLQGLPHLNPFSLHQAQMYNGVGNKV 109
Db 1828 DOQETWRTNFIY---NDVFTWSVADAPGSILYTVQHSPPNNPPTAVLSQMYAGWAGMQF 1884
QY 110 KVLVLGNAFTAGKIIICIPGFAAQNISIA---QATMFPHVIADVRLPIEVLPIEDVR 166
Db 1885 RFIVAGSGVFGGRVAAVPIPG-----IEIGPGLVEVRQFPHVVIDARSLPEVITMFDLR 1939
QY 167 NVLEFH-NDNAPTMRVLVCMLYTPL--RASGSSGTDPPVIAGRVLTCPSPDFSEFLVPP 223
Db 1940 PNMYHPTGDPGLVPLVLVSVYNNLINPFGGSTS-----AIQVTVETRPSEDFEFVMIRAP 1994
QY 224 NVEQTKPFSVPNPLPLNTLSNRVPSLIKSMVSRDHGMVQFQ-----NGRVTLD 274
Db 1995 S--SKTVDSISPAGLLTT-----PVLATGVGNDNRWNGQIVGLQFVPGPGFTCNRRHWNLN 2046
QY 275 QLOGTTPTTASQLCKIRGSVFHANGGNGVNLDELGSPYHAFESP-----APIGFP 326
Db 2047 GSTYGWSSPRFADIDHRKGSASYP-GSNATNVLQFWEYANAGS---AVDNPISQVAPDGF 2102

QY 327 DLGECDHMEASPTTQFNTGDIKQINVKQESAFAPHLGTIOADGLSDVSVNTNMIKLG 386
Db 2103 DMSFVFPNGPIPAAGWVGFAIWSN-----SGAPNVTTVQA-----YELG 2144
QY 387 WVPSPVSDGRHGDVDPWVPIPRYGSTLTLEAAQ-LAPPIYPGPGFEA 429
Db 2145 FAT-----GAPGNLQP-----TTNTSGAQTVAKSIYAVVTGTA 2177
RESULT 8
S64740
genome, polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)
N:Contains: vp60 protein
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate AST/89
C:Date: 12-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
C:Accession: S64740; S46944; S49018; S65012
R:Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.
submitted to the EMBL Data Library, May 1995
A:Description: Genomic organization of rabbit hemorrhagic disease virus determined by
A:Reference number: S64740
A:Accession: S64740
A:Molecule type: genomic RNA
A:Residues: 1-2344 <CAS>
A:Cross-references: EMBL:Z49271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552
A:Experimental source: isolate AST/89
R:Boga, J.; Casais, R.; Martin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra
submitted to the EMBL Data Library, July 1993
A:Description: Molecular cloning, sequence and expression of the capsid protein gene
A:Reference number: S46944
A:Accession: S46944
A:Molecule type: genomic RNA
A:Residues: 1650-2344 <BOG>
A:Cross-references: EMBL:Z24757; NID:g515622; PIDN:CAA80881.1; PID:g515623
A:Experimental source: isolate AST/89
R:Parra, F.; Boga, J.A.; Marin, M.S.; Casais, R.
Virus Res. 27, 219-228, 1993
A:Title: The amino terminal sequence of vp60 from rabbit hemorrhagic disease virus su
A:Reference number: S49018; MUID:93255896; PMID:8488721
A:Accession: S49018
A:Molecule type: genomic RNA
A:Residues: 1650-1796 <PAR>
A:Cross-references: EMBL:Z24757
A:Experimental source: isolate AST/89
A:Accession: S65012
A:Molecule type: protein
A:Residues: 1767-1779; 1875-1877, 'X', 1879-1881; 1936-1938, 'X', 1940-1941 <PAW>
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: blocked amino end; polyprotein
Query Match 10.6%; Score 307.5; DB 2; Length 2344;
Best Local Similarity 26.3%; Pred. No. 4.5e-14;
Matches 149; Conservative 67; Mismatches 212; Indels 139; Gaps 25;
QY 12 MDG---TSAGQLVPANTAE-----PISMEPVAGAATAA-----ATAG-----47
Db 1766 MEKARTAPOGEAAGATTATASVPGTTTDCMDPGCVATTASVVTAEENSSASIAIATAGIGPPQ 1825
QY 48 QVNMIDPWIMNNVQAPQGEFTISPNTPGDILFDLQGLPHLNPFSLHQAQMYNGVGNM 107
Db 1826 QVDOQETWRTNFIY---NDVFTWSVADAPGSILYTVQHSPPNNPPTAVLSQMYAGWAGM 1882
QY 108 KVKVLGNAFTAGKIIICIPGFAAQNISIA---QATMFPHVIADVRLPIEVLPIED 164
Db 1883 QRFIVAGIGVFGGRVAAVPIPG-----IEIGPGLVEVRQFPHVVIDARSLPEVITMFD 1937
QY 165 VRNVLEFH-NDNAPTMRVLVCMLYTPL--RASGSSGTDPPVIAGRVLTCPSPDFSEFLV 221
Db 1938 LRPNMYHPTGDPGLVPLVLVSVYNNLINPFGGSTS-----AIQVTVETRPSEDFEFVMIR 1992
QY 222 PPNVQTKPFSVPNPLPLNTLSNRVPSLIKSMVSRDHGMVQFQ-----NGRVT 272

Db 1993 APS--SKTYVDSISPAQLLT-----PVLTVGVNDNRWNGQIVGLQVPVGGFSTCNHRWN 2044
QY 273 LDQLOQTPTTASQCLCKIRGSVFHANGGNYNLTEL-----DCSPYHAFESP-----APIG 324
Db 2045 LNSTYGVSPRFGDIGHRGSASYP-GNNATNVLQFWYANAGS---AIDNPISQVAPDG 2100
QY 325 FDLGECDDHMEASPTQNTGQVQKQINVKQESAFAPHLGTLQADCLSDSVSNVTNMIK 384
Db 2101 FDMSEVPFNGPGIPAAGWVGFAWNSN-----SGAPNVTVQA-----YE 2142
QY 385 LGWVSPVSDGHRGDVDPWIPRYGSTLTAAQ--LAPPYPPGFGAI-----VFEMSDPFI 439
Db 2143 LGFAT-----GAPGNLP-----TNTSGSQTVAKSIYAVVTGTAQNPAGLFVWASGVI 2191
QY 440 AHGTNGLSVPCPTIQEFTVHFVNEQAPTRGEAALLHYLDPDTHRNLGFEKLYPEGFMTCV 499
Db 2192 STPSANAITYTQPDRIVT-----TPGTAAAPVCKNTPIMFASVV 2232
QY 500 PNSS-----CTGQPTLPI 512
Db 2233 RRTGDVGNATAGSANGTQYCTGSQPLPV 2259

RESULT 9
A:Reference number: A48562
A:Accession: A48562
A:Molecule type: genomic RNA
A:Residues: 1-702 <NE>
A:Cross-references: GB:M87481; NID:G334882; PIDN:AAA16217.1; PID:G334884
A:Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBIP:113565)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.4%; Score 300.5; DB 1; Length 702;
Best Local Similarity 25.2%; Pred. No. 2.6e-14;
Matches 137; Conservative 66; Mismatches 209; Indels 131; Gaps 25;

QY 13 DGTSGAGQLVPEANTAEPTSMPEPVAGATAAATAGOVNMID-PW-IMNNYVQAPQGEFTI 70
Db 154 DGPGGADIVTEEGQTVVQOQPVPAQSALTTLAAASTGKTVDCEWTTFFSYHTA----VNM 209

QY 71 SPNNTPGDILFDLQPLNPLFLSLAOMYNGWGMKVKVLLAGNAFTAGKIIISCIPP 130
Db 210 SITEAQGKLFRRALSPENLPYLRHLSLYTSWSGCIDVRFVTSVSGVGGKLAALIVPP 269

QY 131 GFAAQNISTAQATMFPVHVIADRVLEPIEVPLEDVRNVLFNHNDNAPTMRVLCMLYTP- 189
Db 270 GI--BPVESPTMLQPHVLFDARQTEPVITFDIRKTLYHSMDDTDTRLVIMVYNELI 327

QY 190 -----RASGSSGTDTPFIAGRVLTCPSPDFSLFVPPNVNEQTKPFSVFN--LPLMT 241
Db 328 NPYEOPSEPKSSCSIT-----VETRPSSDFTFSLKPPG--SLLKHGSIPLSDIPRNS 377

QY 242 --LSNRVPSLTKSMVSRDHGQWQFQNGR-----VTLD----- 274
Db 378 RHMGNRWWSITDGEVV-----QPRVQSNRHFDFDSTTTGTWSTPYPIEVTLEKLDRG 432

QY 275 GOLQGTPTTSASQL-----CKIRGSVFHANGGNYNLTEL-----DCSPYHAFESPA 321
Db 433 GOYFKVTDTEKSLVCLPDGWDPTTIPTAMTASNGNYDYTVAEYRTNNGTHFKGFYING 492

QY 322 PI-----GPPDLGECDDHMEASPTQNTGQVQKQINVKQESAFAP-----HLGTQIA 369

Db 493 NLTTKVKGSNDLGET---QOTSRTLTFASVG-----NYKQDQNTINPTHKITSNSLVYDA 543
QY 370 DGLSDSVSVNT---NNIAKLGWV---SPVSDGHRGDVDPWIPRYGVGSLTLEAAQLA--P 419
Db 544 NVVSAATAKTTTWHSTMSHLGYVLVDESVP-----GSDSTKVVRIATLP 587
QY 420 PIYPPGFGAIVFFMSDFPIAHGTNGLSVPCPTIQEFTVHFVNEQAPTRGEAALLHY-LD 478
Db 588 EAFTNG-GNFPVFETTKIOIGH-----FDRAHTKCFNSQVLMTSOKLAENHYTLP 636
QY 479 PDT 481
Db 637 PDS 639

RESULT 10
QY2354
capsid protein - feline calicivirus (strain NADC)
C:Species: feline calicivirus
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2354
R:Seal, B.S.; Bidpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of varia
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2354
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:Cross-references: GB:L09718; NID:G305104; PIDN:AAA16485.1; PID:G305105
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein

Query Match 9.6%; Score 277.5; DB 2; Length 668;
Best Local Similarity 23.3%; Pred. No. 1.3e-12;
Matches 124; Conservative 75; Mismatches 178; Indels 155; Gaps 24;

QY 34 EPVAGAATAATAGOVNMIDPWIMNNYVQAPQGEFTI-----SPNNTPGDILFDLQGLP 87
Db 145 EPSAQMSTAADMATGKSVDSW-----EAFESFHTSVNMSTSETGKILFKQSLGP 195

QY 88 HLNPLSLHIAOMYNGWGMKVKVLLAGNAFTAGKIIISCIPP-FAAQNISTAQATMFP 146
Db 196 LLNPLYLHLSKLYVAMSGVSEVRFSGVFGKLAALIVVPPGVDPVQSTSMLO---YP 252

QY 147 HVIADRVLEPIEVPLEDVRNVLFNHNDNAPTMRVLCMLY---TPLRASGSSGTDPFV 202
Db 253 HVLFDARQVDPVIFSLPDLRSTLYHLMPTDITSLVIMVYNDLINPYANDSNSSGC---I 309

QY 203 IAGRVLTCPSPDFSLFVLP----- 223
Db 310 VT--VETKPGDPFKFHLKPPGSMTHGVSPLDLPKSSSLMIGNRYWSDITDFVVRPFV 367

QY 224 -----NVEQTKTPFSVNL-PLN-TLSNRVPSLTKSMVSRDHGQWQFQNGRVTLD 274
Db 368 FOANRHFNENQETAGWSAPFRPITITISEKSKL--GIGVATDY-----IVP 414

QY 275 GOLQGTPTTSASQLCKIRGSVFHANGGNYNLTELDSGYHAFE-SPAPIGPDLGEC- 332
Db 415 GIPDGMPTTIAEDLTTPAGD-YAITSNGNDIT--TGSEYDSTEVTKNTNFRGMYICGS 471

QY 333 ----WHMEASPTQF-----NTGDVQKQINVKQESAFAPHLGTLQADGLSD 374
Db 472 LQAWCDKKISNTAFITTAKEGNKRPSNTIDMTK-IAYQDT-----HVG-----BE 519

QY 375 VSVNTNMIKLGWVSPVSDGHRGDVDP-----WVIPRYGSLTLEAAQLAPPYPPGFGAI 430
Db 520 VQTSDDALALLGYTGIGEAIGSDRVRVIRSVLPEVGAR-----CGNHP 564

QY 431 VEFMSDFPIAHGTNGLSVPCPTIQEFTVHFVNEQAPTRGEAALLHY-LDPDT 481
Db 565 IFYKSIKLGVIYRSDV-----FNSQILHTRQSLNHYLLSPDS 605

VCWVFC
coat protein - feline calicivirus (strain Japanese F4)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C:Accession: B40481
R:Tohya, Y.; Taniguchi, Y.; Takahashi, E.; Utagawa, E.; Takeda, N.; Miyamura, K.; Yam
Virology 183, 810-814, 1991
A:Title: Sequence analysis of the 3'-end of feline calicivirus genome.
A:Reference number: A40481; MUID:91306470; PMID:1853578
A:Accession: B40481
A:Molecule type: genomic RNA
A:Residues: 1-668 <TOH>
A:Cross-references: GB:D90357; NID:g221264; PIDN:BAAL4371.1; PID:g221266
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predict
F:177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 9.5%; Score 274; DB 1; Length 668;
Best Local Similarity 31.0%; Pred. No. 2.3e-12;
Matches 74; Conservative 39; Mismatches 82; Indels 44; Gaps 9;

QY 23 PEANTA-EPISMEPVAGATAAATAGQVNMIDPMNNVYQAPQGEFTI-----SPNNT 75
DB 133 PEQTAGVGGVIAEPEAQAQMSAADMASCKSVDSW-----EAFSEFTSVNMSSTSET 183
QY 76 PGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVKVLLAGNAFTAGKIIISCIPPGF-AA 134
DB 184 QGKILFQSLGFLNLPYLEHLSKLYVAMWSGIEVRSISGSGVFGKLAIVVPPGVDPV 243
QY 135 QNISTAQATMPPHYADRVLEPIEVPLEDVRNVLFHNNNDNAPTMRVLCMLY---TPLR 190
DB 244 QSTSLQ---YPHVLFDARQVPIFTIPDLRSTLYHVMDSDTTTSVLIMVYNDLINPYA 300
QY 191 ASGSSGCTDPPIAGRVLTCPSPDFSLFLVPPNVEQKTPFSPVNLPLNTLSNSRVPs 249
DB 301 NDSNSSGC---IVT--VETKPGDPFKHLLKPPG-----SVLTHGSIPS 339

RESULT 13
VCWVFF
coat protein - feline calicivirus (strain CFI/68 FIV)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Aug-1999
C:Accession: A40507; B40507; T09246
R:Neill, J.D.; Reardon, I.M.; Heinrikson, R.L.
J. Virol. 65, 5440-5447, 1991
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline cali
A:Reference number: A40507; MUID:91374597; PMID:1716692
A:Accession: A40507
A:Molecule type: genomic RNA
A:Residues: 1-668 <NEI>
A:Cross-references: GB:M32819; NID:g323874; PIDN:AAA2925.1; PID:g323875
A:Accession: B40507
A:Molecule type: protein
A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>
R:Neill, J.D.
submitted to the EMBL Data Library, April 1998
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.
A:Reference number: Z16626
A:Accession: T09246
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: genomic RNA
A:Residues: 1-668 <NE3>
A:Cross-references: EMBL:U13992; NID:g3056875; PIDN:AAC13993.1; PID:g537256
A:Experimental source: strain CFI/68 FIV
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predict
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 9.3%; Score 269; DB 1; Length 668;
Best Local Similarity 31.0%; Pred. No. 5.4e-12;

VCWVFF
coat protein - feline calicivirus (strain F9)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 28-Jul-2000
C:Accession: B43382; C45538; PQ0407; S23702
R:Carter, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.
Virology 190, 443-448, 1992
A:Title: The complete nucleotide sequence of a feline calicivirus.
A:Reference number: A43382; MUID:92410623; PMID:1529544
A:Accession: B43382
A:Molecule type: genomic RNA
A:Residues: 1-671 <CAR>
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879
R:Carter, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.
Arch. Virol. 122, 223-235, 1992
A:Title: Identification and sequence determination of the capsid protein gene of feline
A:Reference number: A45538; MUID:92117861; PMID:1731695
A:Accession: C45538
A:Molecule type: genomic RNA; protein
A:Residues: 1-671 <CAR>
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879
A:Experimental source: strain F9
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)
R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.
J. Gen. Virol. 73, 2429-2433, 1992
A:Title: The cloning, sequencing and expression of a major antigenic region from the fel
A:Reference number: PQ0407; MUID:93019089; PMID:1402818
A:Accession: PQ0407
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>
A:Cross-references: PIDN:AA823553.1; PID:g257083
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.5%; Score 276; DB 1; Length 671;
Best Local Similarity 26.4%; Pred. No. 1.6e-12;
Matches 102; Conservative 54; Mismatches 128; Indels 102; Gaps 17;

QY 13 DGTSGAGQLVPEANT-AEPISEMEPVAGATAAATAGQVNMIDPMNNVYQAPQGEFTI- 70
DB 127 DGSITA----PEQGTVMGGVIAEPEAQAQMSAADMATCKSVDSW-----EAFSEFH 173
QY 71 -----SPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVKVLLAGNAFTAGKIII 125
DB 174 TSVNNSTSETQKILFKOSLGLPLNLPYLEHLAKLYVAMWSGIEVRSISGSGVFGKLA 233
QY 126 SCIPPGF-AAQNISTAQATMPPHYADRVLEPIEVPLEDVRNVLFHNNNDNAPTMRVLCM 184
DB 234 IVVPPGVDPVQSTMLQ---YPHVLFDARQVPIFTIPDLRSTLYHVMDSDTTTSVLIM 290
QY 185 LY-----TPLRASGSSGCTDPPIAGRVLTCPSPDFSLFLVPPNVEQKTPFSPVNPPLN 240
DB 291 VYNDLINPYANDANSSGC---IVT--VETKPGDPFKHLLKPPG-----SMLTHG 335
QY 241 TLSNSRVPSLIKSMVSRDHGQKQVFNQGRVTLDCQLQGTFTPSASQLCKIRGSGVFHANG 300
DB 336 SIPSDLIPKTSLLWIGNYWSDTIDP-----VIRPFVQANR 372
QY 301 GNGYN-----LTELDDGSPYHA-----PESPA-PIGFPDL---GEC--- 331
DB 373 HFDNFQETAGNSPFRPRISVITEONGAKLIGVATDVIYVPGIDGWPDDTIPGELIPA 432
QY 332 -DWHM-----EASPTQFNQGVK 350
DB 433 GDVAITNGTNDITTATGYDTADIIK 458

RESULT 12

Matches	74;	Conservative	38;	Mismatches	83;	Indels	44;	Gaps	9;
Qy	23	PEANT-APISMEPVAGAAATAAGQVNMIDPWIMNMYVQAPQGEFTI-----SPNNT	75						
Db	133	PEQTMVGGVTAEPNAQMSAADMATKGVSDSEW-----EAFSFTSVNMSTSET	183						
Qy	76	PGDILFDLQGLPHLNPFLSHLAQNYNGVGNMKYKVLLAGNAFTAGKIIISCIPPGF-AA	134						
Db	184	QKILFQSLQGLPLNPYLTHLAKLYVAWSGVSDVRFSGSGVGGKLAATVVPDGPDPV	243						
Qy	135	QNISIAQATMPHHVIADRVVLEIEVPLEDVRRVLFHNNDNAPTMRVLCMLY-----T	190						
Db	244	QSTSMLO---YPHVLEDAQVEPVIESIPDLRSLTYHLMSDTSOTSLVIMYNDLINPYA	300						
Qy	191	ASGSSGTDPPVIAAGRVLTCPSPDFSLFVPPNVNVEOKTKPFSVPNLPNLTLSNSR	249						
Db	301	NDSNSGC---IVT---VETKPGDPKPHLLKPPG-----SMLTHGSIPS	339						
RESULT 14									
C48562									
coat protein - San Miguel sea lion virus (serotype 4)									
N:Alternate names: capsid protein									
C:Species: San Miguel sea lion virus									
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999									
C:Accession: C48562									
R:Neill, J.D.									
Virus Res. 24, 211-222, 1992									
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lion virus									
A:Reference number: A48562; MUID: 92410750; PMID: 1529644									
A:Accession: C48562									
A:Molecule type: genomic RNA									
A:Residues: 1-703 <NEI>									
A:Cross-references: GB:M87482; NID:g334886; PIDN:AAA16220.1; PID:g334888									
A:Note: sequence extracted from NCBI backbone (NCBIP:113567)									
C:Superfamily: feline calicivirus coat protein									
C:Keywords: capsid protein; coat protein; glycoprotein									
F:89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted									
Query Match 9.1%; Score 264; DB 1; Length 703;									
Best Local Similarity 27.4%; Pred. No. 1.4e-11;									
Matches 104; Conservative 53; Mismatches 181; Indels 42; Gaps 17;									
Qy	13	DGSGAGQLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNMYVQAPQGEFTISP	72						
Db	154	DGPGSAEIVTEEGTVVQQQPAPAPTALATLATASTGKSVQEHMTFF--SYHTSINWST	211						
Qy	73	NTPGCDILFDLQGLPHLNPFLSHLAQNYNGVGNMKYKVLLAGNAFTAGKIIISCIPPGF	132						
Db	212	VESQCKILYSQLNPSINPYLDHLAKLYSTWSGGIDVRFVTSVSGVGGKLAALLVPPGV	271						
Qy	133	-AONISIAQATMPHHVIADRVVLEIEVPLEDVRRVLFHNNDNAPTMRVLCMLYTPLRA	191						
Db	272	EPIESVSMLO---YPHVLFDAQTEPVIFIPDLIRKTLFGLMSDTSOTSLVIMYNDLINPY	324						
Qy	192	SGSSSGTDPPVIAAGRVLTCPSPDFSLFVPPNVNVEOKTKPFSVFN--LPLNTL--SNSRV	247						
Db	325	NGVENKTTCSI---TVETRPSADFTFALLKPPG--SLIKHGISPDLIPRNSAHWMGNRW	379						
Qy	248	PSLIKSMVSRDHQVMQVFQNGR-VTLDDGLOG-TTPTSASQLCKIRGSVFHANGNG--	303						
Db	380	WSTISGSFV-----QPRVFSQNRHFDSDTTTGWSTPYVPIETKIQGV---GSKNWK	430						
Qy	304	YNLTGDSGPVHAPESPAPITGFGDLGCDWHMEASPTT-OFNTGDVTK--QINVKQESAF	360						
Db	431	FHVIDTD---KALVPGIPDGPWPTTIPD---ETKATNGNFSYGESYRAGSTTIKPNENS	483						
Qy	361	APHLGTTQADGLSDVSVNNT	380						
Db	484	THFKGTVICGTLSTVEIPEN	503						

THIS PAGE BLANK (USPTO)

QY	101	LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFVSPNPLN	240
DB	181	LVCMLYTPLRGSGSGSDSEVWAGRVLTAPSSDFSLFLVPPTEQKTRAFVNPILQ	240
QY	241	TLSNRVPSLKSMVMRDHQVQFQNGRVTLTGQLTGTTPTTSASOLCKIRGSVFHANG	300
DB	241	TLSNRFPFSLQGLLSPDASOVQVQFQNGRCLIDQLLGTTPATSGQLFRVRGKI--NOG	298
QY	301	GNYNLTEDGSPYHAFESPAPIGFPPDLGECDDHWEASPT--TQFNTGDVTKQINVKOE-S	358
DB	299	ARTNLTEVDGKPFNAFSPAPVGFDFGKCDWHMRISKTNNNTGSGDPPHRSVSQVNVQ	358
QY	359	AFAPHLGITQADGLSDVSVNTMIKLGWSPVSDGHRGVDYDPWVPIRYGSLTEFAAOLA	418
DB	359	GFVPHLGSIQFDEVFNHTPG--DYIGTIEWISQSPPTPPGTDINLWEIPDYGSSLSQAANLA	417
QY	419	PIYPFGGEATVFMSPDFPIAHGTNGLS-----VPCITPQEFVHFVNEQAPTRGEALL	474
DB	418	PPVFPFGGEALVYFVSAPP--GPNNRSAPNDVPCLLPQXYITHFVSEQAPTMGDAALL	474
QY	475	HYLDPTDTHNLGFEKLYPGEFMTCPVNSSGTGPQTLPLNGVFEVFSWVSRYQLKPVGTA	534
DB	475	HYVDPTDTHNLGFEKLYPGGTYLTCVPNGVGAGPQQLPLNGVFEVFSWVSRYQLKPVGTA	534
QY	535	GPA-CRLGIRR	544
DB	535	STARGRLGVR	545
RESULT 2			
ID	COAT_LORDV	STANDARD;	PRT; 539 AA.
AC	P54635;		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DT	01-NOV-1997 (Rel. 35, Last annotation update)		
DE	Coat protein (Capsid protein).		
OS	Lordsdale virus (human enteric calicivirus).		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;		
OC	Norovirus.		
OX	NCBI_TaxID=82658;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=96005060; PubMed=7561776;		
RR	Dingle K.E., Lambdon P.R., Caul E.O., Clarke I.N.;		
RT	"Human enteric Caliciviridae: the complete genome sequence and		
RT	expression of virus-like particles from a genetic group II small		
RT	round structured virus."		
RL	J. Gen. Virol. 76:2349-2355(1995).		
CC	-1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; X86557; CAA60255.1;		
DR	InterPro; IPR004005; Calicic_coat.		
DR	Pfam; PF00915; Calicic_coat; 1.		
KW	Coat protein; Glycoprotein.		
FT	CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).		
SQ	SEQUENCE 539 AA; 58775 MW; 5E5C63E7F2C3FD21 CRC64;		
Query Watch			
Best Local Similarity 44.9%; Pred. No. 2.8e-80;			
Matches 254; Conservative 77; Mismatches 186; Indels . 49; Gaps 16;			
QY	1	MMASKDAPTNWDGTSAGQLVPEANTAEPISEMPVAGAAATAAATAGQVNMIDFWIMNNY	60
DB	1	MKMASNDANPS--DGS--AANLVPEVNN--EYWALPEVGAATAAPVAGQONVIDWIRNNE	56
QY	61	VOAPOGEFTISPNNTPGDILFDQLGPHLNPFLSLHAQMYNGWGNKVKVLLAGNAFTA	120
DB	57	VOAPGGEFTVSPNAPGELIWSAPLGPLNLYSLSRMYNGYAGGFEVQVILAGNAFTA	116
QY	121	GKIIISCIIPPGFAAONIATAQATMFPPIADRVLEPVEPDLVRRNVLPH--NNDNAPTM	179
DB	117	GKVIFAAPVNPFPTEGLSPSPMFPPIIVDRQLPEVLPLPDVRRNFYHYNQANDSTL	176
QY	180	RLVCMLYTLPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQTKPFVSPNPL	239
DB	177	KLIALMYTLPLRA--NNAGDDVETVSCRVLTRESPDFDFLFLVPPTEVSRTRKPFVPLVT	234
QY	240	NLTLSNRVPSLKSMVMRDHQVQFQNGRVTLTGQLTGTTPTTSASOLCKIRGSVFHAN	299
DB	235	EEMSNSRFPTEPLEKLYTPGSSAFVQVQNGRCCTDGVLLGTTQLLSAVNICNFRGDYTHA	294
QY	300	GGNGY--NLTELDGSPYHAFES--PAPIGFPDL-----GECDDHWEASPTTQFNT	345
DB	295	GSHDYTNMLASQWNSNYDTEELPAPLGPDPFGVKIQGLLTQTTRADGSTRAHKAT--VST	353
QY	346	GDVTKQINVKQESAFAPHLGTIQ--ADGLSDVSVNTMIKLGWSPVSDG--HRGDVDP	401
DB	354	GSV-----HFTPKLGSVQFTTDTNDFQAGN--TKFTPVGVIOGDGHQNEPQQ	401
QY	402	WVIPRYGSLTEAAQLAPPTYPGFGAEIVFPMDSDFPIAHGTNGLSVPCITPQEFVTHFV	461
DB	402	WSPNYSRGTHNVHLAPAVAPTFPGEQLLFFRSTWPGCGSYENMNLDCLLPQEWVLHEY	461
QY	462	NEQAPTRGEAALLHYLDPDTHRLNGEFLKYPEGFMTCVPMNSSGTGPQ--TLPLNGVYFV	519
DB	462	QEAAPACSDVALLREVPNDTGRVLFECKLHSGYIIV---AHTGPYDLVLPPNGYFRED	517
QY	520	SWVSFRFQKPVGTAGPACRLGIRRS	545
DB	518	SWVNFQYTLAPMGNG-----TGRRRR	538
RESULT 3			
ID	POLN_RHDV	STANDARD;	PRT; 2344 AA.
AC	P27410;		
DT	01-AUG-1992 (Rel. 23, Created)		
DT	01-AUG-1992 (Rel. 23, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Non-structural polyprotein [Contains: RNA-directed RNA polymerase		
DE	(EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like		
DE	protein); Coat protein].		
OS	Rabbit hemorrhagic disease virus (RHDV).		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;		
OC	Lagovirus.		
OX	NCBI_TaxID=11976;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=91361557; PubMed=1840711;		
RR	Meyers G., Wirblich C., Thiel H.-J.;		
RT	"Rabbit hemorrhagic disease virus -"		
RT	sequencing of a calicivirus genome.";		
RL	Virology 184:664-676(1991).		
CC	-1- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE		
CC	FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.		
CC	-1- CATALYTIC ACTIVITY: N nucleoside triphosphate " N diphosphate +		
CC	{RNA}(N).		
CC	-1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC	EMBL; M67473; AAA47285.1; .	OS	San Miguel sea lion virus (serotype 1) (SMSV 1).
CC	PIR; A41039; RRMWRH.	OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
DR	MEROPS; C24.001; .	OX	NCBI_TaxID=36406;
DR	InterPro; IPR004005; Calici_coat.	RN	[1]
DR	InterPro; IPR004004; Calici_pol_hel.	RP	SEQUENCE FROM N.A.
DR	InterPro; IPR000317; Endoptase_C24.	RX	MEDLINE-92410750; PubMed-1529644;
DR	InterPro; IPR000605; RNA_helicase.	RA	Neill J.D.;
DR	InterPro; IPR001205; RNA_pol_P3D.	RT	"Nucleotide sequence of the capsid protein gene of two serotypes of
DR	Pfam; PF00680; RNA_dep_RNA_pol; 1.	RT	San Miguel sea lion virus: Identification of conserved and non-
DR	Pfam; PF00910; RNA_helicase; 1.	RT	conserved amino acid sequences among calicivirus capsid proteins.";
DR	Pfam; PF00915; Calici_coat; 1.	RL	Virus Res. 24:211-222(1992).
DR	Pfam; PF03510; Peptidase_C24; 1.	CC	-!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
DR	PRINTS; PR00916; 2CENDOPTASE.	CC	-----
DR	PRINTS; PR00918; CALICIVIRUSN.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
KW	Polypeptide; Transferase; RNA-directed RNA polymerase; Hydrolase;	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
FT	CHAIN 7508 7707 HELICASE (P2C-LIKE).	CC	the European Bioinformatics Institute. There are no restrictions on its
FT	CHAIN 1109 1251 PROTEASE P3C.	CC	use by non-profit institutions as long as its content is in no way
FT	CHAIN 21497 21625 RNA-DIRECTED RNA POLYMERASE.	CC	modified and this statement is not removed. Usage by and for commercial
FT	CHAIN ? 2344 COAT PROTEIN.	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
FT	NP_BIND 522 529 ATP (POTENTIAL).	CC	or send an email to license@isb-sib.ch).
FT	ACT_SITE 1212 1212 PROTEASE (POTENTIAL).	CC	-----
FT	ACT_SITE 1227 1227 PROTEASE (POTENTIAL).	DR	EMBL; M87481; AAA16217.1; .
FT	SEQUENCE 2344 AA; 257066 MW; 1454C24F81E9212 CRC64;	DR	PIR; A48562; A48562.
Query Match	10.9%; Score 316; DB 1; Length 2344;	DR	InterPro; IPR004005; Calici_coat.
Best Local Similarity	28.3%; Pred. No. 8.8e-16;	DR	Pfam; PF00915; Calici_coat; 1.
Matches 132; Conservative 60; Mismatches 172; Indels 102; Gaps 21;		KW	Coat protein; Glycoprotein.
QY	12 MDGTSAGQLVPEANTAEPIIS-----MEPVAGAAATAA-----ATAG-----47	FT	CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
Db	1766 MEGKARAAPOGEAGTAATTASPGTTDCMDPGVVAITSVITAENSASSIATAGIGPPQ 1825	FT	CARBOHYD 481 481 N-LINKED (GLCNAC. .) (POTENTIAL).
QY	48 QVNMIDPWIMNMYVOAQGEFTISNPNTPGDLFDLQGLHNPFLSLHQAQYNGWGM 107	FT	CARBOHYD 493 493 N-LINKED (GLCNAC. .) (POTENTIAL).
Db	1826 QVDQETWRTNFY---NDVFTWSVADAPGSLTYVQHSPPNNPETAQLSQMYAGWAGM 1882	FT	CARBOHYD 545 545 N-LINKED (GLCNAC. .) (POTENTIAL).
QY	108 KVKVLLAGNAFTAGKIISCIPPFGAAQNIATA---QATMEPHVIADVRLPIEVPLED 164	SEQ	SEQUENCE 702 AA; 77850 MW; E6E5A58523DEE3D7 CRC64;
Db	1883 QRFVIVAGSGVFGGLVRAVIPP-----IEIGPGLVQRQPPHVVIDARSLPEVTTMPD 1937	Query Match	10.4%; Score 300.5; DB 1; Length 702;
QY	165 VNRVLFH--NNDNAPTMRLVCMLYTPL--RASGSSSGTDPFVIAGRVLTCPSPDFSF 221	Best Local Similarity	25.2%; Pred. No. 2.6e-15;
Db	1938 LRPNMVHPTGDPGLVPTLVSYNNLINPFGGTS-----AQVIVETRPSEDFEVMIR 1992	Matches 137; Conservative 66; Mismatches 209; Indels 131; Gaps 25;	
QY	222 PPNVQKTKPFVSVNPLPLNTLSNRPVSLIKSMVSRDHQGMVQFQ-----NGRVT 272	QY	13 DGTSGAGQLVPEANTAEPIISMPEVAGAAATAATAAGOVNMID-PW-IMNMYVOAQGEFTI 70
Db	1993 APS--SKTVDSISPAGLTT-----PVLTVGNDNRWNGQIVGLQVPVPGGFSTCNRHWN 2044	Db	154 DCPGGADIVTEEGQIVVQOQPVPAOSALTTLAAASGKTVDCEWTTFFSYHTA----VNM 209
QY	273 LDQLOGTTPTSASOLCKIRGSVFHANGNGYNLTET---DGSPYHAFESP-----APTG 324	QY	71 SPNPTPGDILFDLQGLHNPFLSLHQAQYNGWGMVKVLLAGNAFTAGKIISCIPP 130
Db	2045 LMGSTYGVSSPFGDIDHRGSAST--SGSNATNVLQFWYANAGS---ALDNPISQVAPDG 2100	Db	210 SSTEAGQKILFSLRALSPELNPLYLRHISLSYTWSSGIDVRFVTSVSGVFGGKLAALIVPP 269
QY	325 FPDLGECMDHMEASPTQFNTGDVIKQINVKOESAFAPHLGTLQADGLSDSVNTNIAK 384	QY	131 GFAAONISIAQATMPHVIADVRLPIEVPLEDVRNVLFNHNDNAPTMRLVCMLYTPL- 189
Db	2101 PDMSEVPENGPPIAAGWVGFAIWSN-----SGAPNVTVQV-----YE 2142	Db	270 GI--EPVESPTMLQYPHVLFVDARQTEPVITIPDIRKTLYHSMDDDTDTTRLVIMVYNELI 327
QY	385 LGWSPVSDGHRGVDVPPVIRYGGTTLTAAQ--LAPPIYPPGFGEA 429	QY	190 -----RASGSSSGTDPFVIAGRVLTCPSPDFSFVLVPPNVEQKTKPFVSPN--LPLNT 241
Db	2143 LGFAT----GAPGNLQ-----TTNTSGAQTVAKSIVAVVTGTA 2177	Db	328 NPYEQSEPKSSCSIT-----VETRPSSDTFSLKKPPG--SLLKHGSIPLDLIPRNS 377
RESULT 4		QY	242 ---LSNRPVSLTKSMVSRDHQGMVQFQNGR-----VTLD-----274
COAT_SMSV1	STANDARD; PRT; 702 AA.	Db	378 RHWGMNRWMTIDGFVV-----QPRVFQSNRHFDEFTTGTWSTPYPIEIVTELEKLRG 432
AC	P36284;	QY	275 GOLQGTTPTSASOL-----CKIRGSVFHANGNGYNLTET---DGSPYHAFESPA 321
DT	01-JUN-1994 (Rel. 29, Created)	Db	433 GQYFKVTDTEKSLVPLGDPGWPDTTIPTAMTASNGNYDYTVAEYRITNNGTHFKGYIMG 492
DT	01-JUN-1994 (Rel. 29, Last sequence update)	QY	322 PT-----GFPDLGECMDHMEASPTQFNTGDVIKQINVKOESAFAP-----HLCTIOA 369
DT	15-DEC-1998 (Rel. 37, Last annotation update)	Db	493 NLTTKVKGSDNLGET--QOTSRTLFASVG-----NYKQNTINPTHTKITSNSLVVYDA 543
DE	Coat protein (Capsid protein).	QY	370 DGLSDSVNVT-----NMIKLGWV-----SPVSDGHRGVDVPPVIRYGGTTLTAAQIA--P 419
		Db	544 NNVSAAATAKTTTWHSTMSHLGYVLVDESVP-----GSDSTKVVRITATLP 587
		QY	420 PIYPGFGCAIYFFMSDFPIAHGTNGLSVPCIPIQGFVTHFVNEQAPTMGEAALLHY-LID 478
		Db	588 EAFENG-GNFPVFFFTNKIQIGH-----FDRAHTKCFNSQVLMTSQKLAENHYTLP 636

Thu Jan 16 15:59:23 2003

```
QY 479 PDT 481
Db 637 PDS 639

RESULT 5
POLN_MANCV STANDARD: PRT: 2208 AA.
AC Q69014;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: RNA-directed RNA polymerase
DE (EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like
DE protein); Coat protein].
OS Manchester virus (Human enteric calicivirus).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Sapovirus.
OX NCBI_TaxID=82659;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=95390791; PubMed=7661689;
RA Liu B.L., Clarke I.N., Caul E.O., Lambden P.R.;
RT "Human enteric caliciviruses have a unique genome structure and are
RT distinct from the Norwalk-like viruses.";
RL Arch. Virol. 140:1345-1356(1995).
CC -1- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).
CC -1- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR
CC THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X86560; CAA60262.1; --
DR MEROPS; C24.UPW; --
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR000317; Endoptase_C24.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR Pfam; PF00915; Calici_coat; 1.
DR Pfam; PF03510; Peptidase_C24; 1.
DR PRINTS; PR00916; 2CENDOPTASE.
DR PRINTS; PR00918; CALICIVIRUSN.
DR Polyprotein; transferase: RNA-directed RNA polymerase; Hydrolase;
KW Thiol protease; Helicase; ATP-binding; Coat protein.
FT CHAIN ? ? HELICASE (P2C-LIKE).
FT CHAIN ? ? PROTEASE (P3C).
FT CHAIN ? ? RNA-DIRECTED RNA POLYMERASE.
FT CHAIN ? 2208 COAT PROTEIN.
FT NP_BIND 408 415 ATP (POTENTIAL).
FT ACT_SITE 1097 1097 PROTEASE (BY SIMILARITY).
FT ACT_SITE 1112 1112 PROTEASE (BY SIMILARITY).
SQ SEQUENCE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;

Query Match 10.0%; Score 289; DB 1; Length 2208;
Best local similarity 22.9%; Pred. No. 9.4e-14;
Matches 138; Conservative 82; Mismatches 244; Indels 138; Gaps 25;

14 GTSAGQLVPEANTAEPISEPVAGAA----TAAATAGQVNMIDPMNMNVQAPQGEFT, 69
OY
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M86379; AA79327.1; -
 DR EMBL; Z11536; CA77636.1; -
 DR PIR; B43382; VCMWFC.
 DR PIR; C45538; C45538.
 DR InterPro: IPR004005; Calici_coat.
 DR Pfam: PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 618 618 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 671 AA; 73441 MW; 33BEE86D8370D5E5 CRC64;

Query Match 9.5%; Score 276; DB 1; Length 671;
 Best Local Similarity 26.4%; Pred. No. 1.8e-13;
 Matches 102; Conservative 54; Mismatches 128; Indels 102; Gaps 17;

QY 13 DFTSGAGQLVPEANT-AEPISEMPVAGATAAATAGQVNMIDPWIMNNYVQAPQGEFTI- 70
 DB 127 DGSITA---PEQGTWGVGVIAEPSAQMSTADMATGKSVDSW-----EAFESFH 173
 QY 71 -----SPNTPGDLPLGPHLPFLSHLAQYNGWGNMKVLLAGNAPTAAGKIII 125
 DB 174 TSVNWTSETQGRILFKQSLGPLNLYLEHLAKLYVWAGSIEVRSISGSGVFGKLA 233
 QY 126 SCIPPCF-AAQNTSIAQATMFHIVADVRLVLEPIEDVRNVLPHNDNAPTMLVCM 184
 DB 234 IVVPGVDVQVSTMLQ---YHVLFDARQVEPIETDLRLSTLYHMSDITDTTSLVIM 290
 QY 185 LY----TPLRASGSGTDFVIAGRVLCPSDFSLFLVPPNVQKTKPFSVNPPLN 240
 DB 291 VYNDLINPVANDANSSG---IVT--VETKPGDFKHLKPPG-----SMLTHG 335
 QY -241 TLSNRPVSLIKSMVSRDHQWQVQNGRVLDGLOQTTPTSASQLCKIRGVSFPHANG 300
 DB 336 SIPSDLIPKTSLLWIGNYWSIDTF-----VIRPFVQANR 372
 QY 301 GNGYN-----LTELDSGPYHA-----PESPA-PIGEFDL---GEC--- 331
 DB 373 HFDNQETAGWTPRPRIISVITEQNGAKLGIGVATDYIVQIPDGPWPTTIPGELIPA 432
 QY 332 -DWHM-----EASPTQFNTGDKV 350
 DB 433 GDVAITNGTGNDITATGYDTADIIK 458

RESULT 7
 COAT_FCVF4
 ID COAT_FCVF4 STANDARD; PRT; 668 AA.
 AC P27405;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Coat protein (Capsid protein).
 OS Feline calicivirus (strain Japanese F4) (FCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OC NCBI_TaxID=11980;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91306470; PubMed=1853578;
 RA Tohya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,
 RA Miyamura K., Yamazaki S., Mikami T.;
 RT "Sequence analysis of the 3'-end of feline calicivirus genome."

RL Virology 183:810-814(1991).
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D90357; BA14371.1; -
 DR PIR; B40481; VCMWFC.
 DR InterPro: IPR004005; Calici_coat.
 DR Pfam: PF00915; Calici_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 668 AA; 73589 MW; 85BDBC85804E503 CRC64;

Query Match 9.5%; Score 274; DB 1; Length 668;
 Best Local Similarity 31.0%; Pred. No. 2.6e-13;
 Matches 74; Conservative 39; Mismatches 82; Indels 44; Gaps 9;

QY 23 PEANTA-EPISMEPVAGATAAATAGQVNMIDPWIMNNYVQAPQGEFTI-----SPNNT 75
 DB 133 PEOGTAVGVVIAEPSAQMSTADMATGKSVDSW-----EAFSPHTSYNNWSTSET 183
 QY 76 PGDILFDLQGLPHLPFLSHLAQYNGWGNMKVLLAGNAPTAAGKIIISCIIPPG-AA 134
 DB 184 QGKILEKQSLGPLNLYLEHLKLYVWAGSIEVRSISGSGVFGKLAIVVPGVDPV 243
 QY 135 QNTSIAQATMFHIVADVRLVLEPIEDVRNVLPHNDNAPTMLVCLMY----TPLR 190
 DB 244 QSTSMQLQ---YHVLFDARQVEPIETDLRLSTLYHMSDITDTTSLVIMVNDLINPYA 300
 QY 191 ASGSSSGTDFVIAGRVLCPSDFSLFLVPPNVQKTKPFSVNPPLNTLSNSRVP 249
 DB 301 NDSNSSG---IVT--VETKPGDFKHLKPPG-----SVLTHGSIPS 339

RESULT 8
 COAT_FCVC6
 ID COAT_FCVC6 STANDARD; PRT; 668 AA.
 AC P27404;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Coat protein (Capsid protein).
 OS Feline calicivirus (strain CFI/68 FIV) (FCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 OC Vesivirus.
 OC NCBI_TaxID=11979;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=91374597; PubMed=1716692;
 RA Neill J.D., Reardon I.M., Heinrikson R.L.;
 RT "Nucleotide sequence and expression of the capsid protein gene of
 RT feline calicivirus."
 RL J. Virol. 65:5440-5447(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Neill J.D.;
 RA Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- PTM: The N-terminus is blocked.
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.

CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

DR PIR; C48562; C48562.

CC -I- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,

```
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -!- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; D00625; BAA00516.1; ALT_SEQ.  
DR PIR; A34032; GNNY2W.  
DR HSP; P03299; IPOV.  
DR MEROPS; C03.001; -.  
DR MEROPS; C03.020; -.  
DR InterPro; IPR000199; Cys-protease-3C.  
DR InterPro; IPR003138; Pico_P1A.  
DR InterPro; IPR000081; Pico_P2A.  
DR InterPro; IPR002527; Pico_P2B.  
DR InterPro; IPR000605; RNA_helicase.  
DR InterPro; IPR001205; RNA_pol_P3D.  
DR InterPro; IPR001676; RNV.  
DR Pfam; PF000073; rhv; 3.  
DR Pfam; PF00548; Cys-protease-3C; 1.  
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.  
DR Pfam; PF00910; RNA_helicase; 1.  
DR Pfam; PF00947; Pico_P2A; 1.  
DR Pfam; PF01552; Pico_P2B; 1.  
DR Pfam; PF02226; Pico_P1A; 1.  
DR Pfam; PF02252; Pico_P2A; 1.  
DR ProDom; PD001125; Cys-protease-3C; 1.  
DR ProDom; PD001274; Pico_P2B; 1.  
DR ProDom; PD001306; Pico_P2A; 1.  
KW Polypeptide; Coat protein; Core protein; Transferrase;  
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.  
FT CHAIN 2 69  
FT CHAIN 70 340  
FT CHAIN 341 578  
FT CHAIN 579 879  
FT CHAIN 880 1028  
FT CHAIN 1029 1125  
FT CHAIN 1126 1454  
FT CHAIN 1455 1541  
FT CHAIN 1542 1563  
FT CHAIN 1564 1746  
FT CHAIN 1747 2205  
FT LIPID 2 2  
FT ACT_SITE 1710 1710  
FT ACT_SITE 1724 1724  
FT ACT_SITE 1724 1724  
SQ SEQUENCE 2205 AA; 245701 MW; 2442AB039E0254AD CRC64;  
  
Query Match 5.1%; Score 146.5; DB 1; Length 2205;  
Best Local Similarity 19.2%; Pred. No. 0.0073;  
Matches 109; Conservative 71; Mismatches 188; Indels 199; Gaps 26;  
  
QY 22 VPEANTAEPISEPVAGATAATAAGVNMIDPTMNNVVOAQPOQEFTISPN-----NTP 76  
DB 365 IFEFDVTPID---IPGEVRNMELAEIDTMIPLNTSORKNTMDMYVELNDAHSDFP 421  
QY 77 GDILFDLQLGPHLNPFLSH-----LAOMYGVGNMVKVLLAGNAFTAGKTIISCIPPG 131  
DB 422 ---ILCLSLSPASDPRLAHTMLGEILNYTHWAGSLKFTFLFCGSMATGKLLVYAPPG 478  
QY 132 FRAQNISIAQATMFHVIAADVRLVLEPIEVLPEEDVRNVLFHN---NNAP-----TMR 180  
DB 479 AKAPE-SRKEAMLGTHVINDIGLQSCTMVVPWISNTTYRQTINDSFTEGGYISMFYQTR 537  
QY 181 LVCMLYTPLRASGSSGTDPPFIAGRVLTCPSPDSEFLF-----VPPN----- 224  
DB 538 VVVLSTPRKMD-----ILGFVSAC---NDFSRLLRDTHTHISQEVMPQGLDLIE 585
```

```
QY 225 --VEQKTKPFSVPNPLPLNTLSNR-----VPS-----LJKSMVSRDHQ 262  
DB 586 GVVEGVTRNALPTLPTVNNLPDRSSGPAHSKETPALTAVETGATNPLVPSDTVQTRH-- 643  
QY 263 MVQ-----PQGRVTLGQLOQTTP--SASQCKKIRGSVP----- 296  
DB 644 VTOKRTRSESTVESFARGACVAIEVDNDAPTTRASKLFSVNMKITYKDTVQLRRKLEFF 703  
QY 297 -----HANGGNGYNLTEDGSPYHAFESPAPIGRPDLG---ECDHMEEA 337  
DB 704 TYSREDMEFTFVVTSNYTDANNGHALNVOYQIMY-----IPGAPIPCKRNDYTWTQTS 757  
QY 338 SPTTQNTGDVQIKQINVKQESAFAPHLGTQA-----DGLSDVSVNTNMIKLGWSPVS 392  
DB 758 NPSVFYTCABPARISV-----PYVGIANAYSHFYDGFAY----- 793  
QY 393 DGHGDDVDPIVPRYGSTLTAAQLAPPIYPCFGCAIVFFMSDFPIAHGTNGLSVPTI 452  
DB 794 -----PLAQASTEGDSL-----YGAA-----SLNDF-----GSLAV----- 820  
QY 453 PQEFVTHFVNEQAPTRGEAALLHYLDP 479  
DB 821 -----RVVNDHNPTKLTSKIRVYMKP 841  
  
RESULT 11  
POLG_POL2L  
ID POLG_POL2L STANDARD; PRT: 2207 AA.  
AC P06210;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins  
DE P2A TO P2C; P3A; Genome-linked protein VP3; Picornain 3C  
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D  
DE (EC 2.7.7.48)]  
OS Poliovirus type 2 (strain Lansing).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Enterovirus.  
OX NCBI_TaxID=12084;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86115399; PubMed=3003384;  
RA la Monica N., Merlam C., Racaniello V.R.;  
RT "Mapping of sequences required for mouse neurovirulence of poliovirus  
RT type 2 Lansing".  
RL J. Virol. 57:515-525(1986).  
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.  
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-Gly bond in the  
CC poliovirus polypeptide. In other picornavirus reactions Glu may be  
CC substituted for Gln, and Ser or Thr for Gly.  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC [RNA](N).  
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -!- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; M12197; AAA46912.1; -.  
DR PIR; A29507; GNNY5P.  
DR HSP; P03299; IPOV.  
DR MEROPS; C03.001; -.  
DR MEROPS; C03.020; -.  
DR
```

DR	InterPro: IPR000199; Cys-protease-3C.	
DR	InterPro: IPR003138; Pico_P1A.	
DR	InterPro: IPR000081; Pico_P2A.	
DR	InterPro: IPR002527; Pico_P2B.	
DR	InterPro: IPR000605; RNA_helicase.	
DR	InterPro: IPR001205; RNA_pol_P3D.	
DR	InterPro: IPR001676; Rhv.	
DR	Pfam: PF00073; rhv. 3.	
DR	Pfam: PF00548; Cys-protease-3C; 1.	
DR	Pfam: PF00680; RNA_dep_RNA_pol; 1.	
DR	Pfam: PF00910; RNA_helicase; 1.	
DR	Pfam: PF00947; Pico_P2A; 1.	
DR	Pfam: PF01552; Pico_P2B; 1.	
DR	Pfam: PF02226; Pico_P1A; 1.	
DR	ProDom: PD001125; Cys-protease-3C; 1.	
DR	ProDom: PD001274; Pico_P2B; 1.	
DR	ProDom: PD001306; Pico_P2A; 1.	
KW	Polyprotein; Coat protein; Core protein; Transferase;	
KW	RNA-directed RNA polymerase; Hydrolyase; Thiol protease; Myristate.	
FT	CHAIN 2 69	
FT	CHAIN 70 340	
FT	CHAIN 341 578	
FT	CHAIN 579 879	
FT	CHAIN 880 1028	
FT	CHAIN 1029 1125	
FT	CHAIN 1126 1454	
FT	CHAIN 1455 1541	
FT	CHAIN 1542 1563	
FT	CHAIN 1564 1746	
FT	CHAIN 1747 2207	
FT	CHAIN 2 2	
FT	LIPID	
FT	ACT_SITE 1710 1710	
FT	ACT_SITE 1724 1724	
FT	SEQUENCE 2207 AA; 245829 MW; 2B1E2070B7D44F99 CRC64;	
SQ		

	Query Match	5.0%;	Score 145.5;	DB 1;	Length 2207;
	Best Local Similarity	18.8%;	Pred. No. 0.0087;		
	Matches	106;	Conservative 72;	Mismatches 193;	Indels 193; Gaps 24;
QY	22 VPEANTAEPISMPEVAGATAAATACGVNMIDPDIMINNYVOAQGEFTSPN-----NTP 76	: :	:	:	:
Dd	365 IPEDEVTPPID----IPGVRNNMELAEIDTPIPLNTNQRKNTMDMYRVELNDAAHSDDP 421	: :	:	:	:
QY	77 GDILFDLQLGHLPNLFSH-----LAQMNGWVGWMKVLLAGNACTACKIIISCIPTPG 131	: :	:	:	:
Dd	422 ---ILCLSLSPASDPRLATMLCEILNIYYTHWAGSKETFLFCGSMMATCKLLIVSTAPGC 478	: :	:	:	:
QY	132 FAQNISIATAWFPHVIADVRLIEPVLEDVDRNVLFHN--NDNPAN-----TMR 180	: :	:	:	:
Dd	479 AEAPRK-SRKEAMLGTHVIDIGLOSCTMVVPWISNTTYROTINDSFTEGGYSIMEFYQTR 537	: :	:	:	:
QY	181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSELF-----LVP 222	: :	:	:	:
Dd	538 VVVPLSITPKMD-----ILGFVSAC---NDFSVRLLRDTTTHISOEAMPQGLGLDIE 585	: :	:	:	:
QY	223 PNVEOKTKPPSVPNLPNTLSNR-----VPS-----LKSMVMYSRDDHQ 262	: :	:	:	:
Dd	586 GVVECVGTVRNALPTLPANNLPDTQSSGPAHSKETALTAVETGATNPLVPSDTVQTRH-- 643	: :	:	:	:
QY	263 MVO-----FONGRVTLDGLOGQTPTS-ASOLCKIRGSVF-----296	: :	:	:	:
Dd	644 YIQKTRSESTVESFFARGACVALIENVDNADPTRKRASKLFSWKITYKTDLVKLRKKLEFF 703	: :	:	:	:
QY	297 -----HANGNGYNLTDELDSGYAHAFESPAPIGPDPDLCGDWHWEASPT 340	: :	:	:	:
Dd	704 TYSRFDEMTTFVTNSVNTDANHALLNQVYQIMYPFGAPIPCKNWDY---TWOTSSNPS 760	: :	:	:	:
QY	341 TFQNTGCDVIKQINKQESAFPHLGTQA-----DGLSDVSVNTNMIAKLGWVSPVSDGH 395	: :	:	:	:
Dd	761 VFETYTGAPPARISV-----PYVGIANAYSHFYDFGAKV-----793	: :	:	:	:
QY	396 RGDVDWPVIRYRGSTLTLEAAQLAPPYPPGFEAIYFWMSDFPIAHGTNGLSVCPITQE 455	: :	:	:	:

794 -----PLAGQASTEGDSL-----YGAA-----SLNDF-----GSLAV----- 820

QY 456 FVTHFVNEQAPTRGEALLHYLDP 479
 ||: ||: : : ||: |
 DB 821 ---RVVNDHNPTKLTSKIRVIMKP 841

RESULT 12
 POLG_POL3L STANDARD; PRT: 2206 AA.
 AC ID POLG_POL3L Q84783; Q84784; Q84785; Q84786; Q84787; Q84788; Q84789;
 AC P03302; Q84790; Q85592; Q85593; Q85594;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 15-JUN-2002 (Rel. 01, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
 DE p2A TO p2C, p3A; Genome-linked protein VPg; Picornain 3c
 DE (EC 3.4.22.28) (protease 3C) (P3C); RNA-directed RNA polymerase P3D
 DE (EC 2.7.7.48)]
 OS Poliovirus type 3 (strains P3/Leon/37 and P3/Leon 12A[1]B).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.
 OX NCBI_TaxID=12088;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=P3/Leon/37; PubMed=6324200;
 RX MEDLINE=84170338; PubMed=6310508;
 RA Stanway G., Hughes P.J., Mountford R.C., Reeve P., Minor P.D.,
 RA Schild G.C., Almond J.W.;
 RA "Comparison of the complete nucleotide sequences of the genomes of
 RT the neurovirulent poliovirus P3/Leon/37 and its attenuated Sabin
 RT vaccine derivative P3/Leon 12alb.";
 RT Proc. Natl. Acad. Sci. U.S.A. 81:1539-1543(1984).
 RL [2]
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=P3/Leon 12A[1]B;
 RX MEDLINE=83299239; PubMed=6310508;
 RA Stanway G., Cann A.J., Hauptmann R., Hughes P.J., Clarke L.D.,
 RA Mountford R.C., Minor P.D., Schild G.C., Almond J.W.;
 RA "The nucleotide sequence of poliovirus type 3 Leon 12 alb: comparison
 RT with poliovirus type 1.";
 RL Nucleic Acids Res. 11:5629-5643(1983).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 1-878.
 RX MEDLINE=95120467; PubMed=7820548;
 RA Grant R.A., Hirameth C.N., Filman D.J., Syed R., Andries K.,
 RA Hogle J.M.;
 RA "Structures of poliovirus complexes with anti-viral drugs:
 RT Implications for viral stability and drug design.";
 RL Curr. Biol. 4:784-797(1994).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1-878.
 RX Hirameth C.N., Grant R.A., Filman D.J., Hogle J.M.;
 RA "Binding of the antiviral drug wins1711 to the Sabin strain of type-3
 RT poliovirus - structural comparison with drug-binding in rhinovirus-
 RL 14.";
 RN [4]
 RP Acta Crystallogr. D 51:473-489(1995).
 CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
 CC O/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-1-Gly bond in the
 CC poliovirus polypeptide. In other picornavirus reactions Glu may be
 CC substituted for Gln, and Ser or Thr for Gly.
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate + N diphosphate +
 CC {RNA}(N).
 CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -!- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -!- MISCELLANEOUS: THE SEQUENCE OF STRAIN SABIN VACCINE P3/LEON/37 IS
 CC SHOWN.
 CC -!- MISCELLANEOUS: THE STRAIN SABIN VACCINE P3/LEON/37 IS THE
 CC PROGENITOR OF THE STRAIN SABIN VACCINE P3/LEON12A[1]B.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

[illegible]

```
FT CHAIN 579 878 COAT PROTEIN VP1.
FT CHAIN 879 1027 CORE PROTEIN P2A.
FT CHAIN 1028 1124 CORE PROTEIN P2B.
FT CHAIN 1125 1453 CORE PROTEIN P2C.
FT CHAIN 1454 1540 CORE PROTEIN P3A.
FT CHAIN 1541 1562 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1563 1745 PICORNAIN 3C.
FT CHAIN 1746 2206 RNA-DIRECTED RNA POLYMERASE P3D.
FT LIPID 2 2 MYRISTATE (BY SIMILARITY).
FT ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
FT ACT_SITE 1723 1723 PROTEASE (POTENTIAL).
SQ SEQUENCE 2206 AA; 245731 MW; F226AD85403C37BA CRC64;

Query Match 4.8%; Score 139.5; DB 1; Length 2206;
Best Local Similarity 20.8%; Pred. No. 0.025;
Matches 96; Conservative 64; Mismatches 182; Indels 119; Gaps 21;

QY 22 VPEANTAEPISEMPVAGATAAATAGOVNMIDPWIMNNVYVQAPQGEFTI---SPNPTPGD 78
DB 365 IPEFDVTPPID---IPGEVKNMELAEIDTMIPLNLENTKRTMDMYRVRLSDSANLSGP 421
QY 79 ILFDLQLGPHLNPFLSH-----LAQMYNGWGNMKVKVLLAGNAFTAGKIIISCIPPGFA 133
DB 422 ILC-LSLSPAADPRLSHTMLGEVLYNYTHWAGSLKFTFLFCGMMATGKLLVAYAPPG-A 479
QY 134 AQNISIAQATMPHHVIADVRVLEPIEVLPLEYVRLFH--NNDNAPTRMLVCMLY----- 186
DB 480 QPPTSKEAMLTGHTVIMDLGLQSSCTMVVPMISNVTYRQTQDSFTGEGYISMFYQTRIV 539
QY 187 TPLRAGSSSGTDPFVIAGRVLTCTSPDFSLFL-----VPPNVE-----QKT 229
DB 540 VPL-----STPKAMDMLGFSVAC--NDFSRLRLDTHTHISOAAMPQGVDDDLITEVAQNA 591
QY 230 KPFSSVP-----NLP-----LNTLS-----NSRVPS 249
DB 592 LALSPLKPOSNLPDTHKASCPAHSKEVPLTAVETGATNPLVSDTVQVRHVITQQRSSS 651
QY 250 LKSM-----MYSRDHQ---MVQ---FONGRYT-----LDQQLGTTTPTSASQLCK 290
DB 652 TIESFARGACVAITEVDNEQATNVQKLFATWRITYKDTVQLRRKLEFTYSRFDMEFT 711
QY 291 IRGSVFHANGNGYNLTLDGSPYHAFESPAPGPDLCGDWHMEASPTTFNGDGVK 350
DB 712 FVVTANTNSNGHALNQVIQIMYPGAPTCKSWDDY---TWQTSNPSIFITYGAAPA 768
QY 351 QINVAQESAFAPHLGTHQIA-----DGLSDSVYNTNNIAKLG 386
DB 769 RLVS-----PYVGLANAYSHFYDGFAPKPLKSANDQVG 802

RESULT 15
POLG_POLIS STANDARD: PRT: 2209 AA
AC Q03301; Q84881; Q84882; Q84883; Q84884; Q84885; Q84886; Q84887;
AC Q84888; Q84889; Q84890;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; P3A; Genome-linked protein VPG; Picornain 3C
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D
DE (EC 2.7.7.48)]
OS Poliovirus type 1 (strain Sabin).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12082;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83299876; PubMed=6310545;
RA Nomoto A., Omata T., Toyoda H., Kuge S., Horie H., Kataoka Y.,
RA Genba Y., Nakano Y., Imura N.;
RT "Complete nucleotide sequence of the attenuated poliovirus Sabin 1
strain genome.";
```

```
RL PROC. Natl. Acad. Sci. U.S.A. 79:5793-5797(1982).
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- MISCELLANEOUS: THIS VIRUS IS A LIVE VACCINE STRAIN DERIVED FROM
CC THE MAHONEY STRAIN BY SPONTANEOUS MUTATIONS DURING THE ATTENUATION
CC PROCESS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V01150; CAA24465.1; -.
CC PIR; A03899; GNNX3P.
CC MEROPS; C03.001; -.
CC MEROPS; C03.020; -.
CC InterPro; IPR000199; Cys-protease-3C.
CC InterPro; IPR003138; Pico_P1A.
CC InterPro; IPR000081; Pico_P2A.
CC InterPro; IPR002527; Pico_P2B.
CC InterPro; IPR000605; RNA_helicase.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR001676; Rhv.
CC Pfam; PF000073; rhv; 3.
CC Pfam; PF00548; Cys-protease-3C; 1.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC Pfam; PF00947; Pico_P2A; 1.
CC Pfam; PF01552; Pico_P2B; 1.
CC Pfam; PF02226; Pico_P1A; 1.
CC ProDom; PD001125; Cys-protease-3C; 1.
CC ProDom; PD001274; Pico_P2B; 1.
CC ProDom; PD001306; Pico_P2A; 1.
KW Polyprotein; Coat protein; Core protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 341 COAT PROTEIN VP3.
FT CHAIN 342 579 COAT PROTEIN VP1.
FT CHAIN 580 881 COAT PROTEIN P2A.
FT CHAIN 882 1030 CORE PROTEIN P2B.
FT CHAIN 1031 1127 CORE PROTEIN P2C.
FT CHAIN 1128 1456 CORE PROTEIN P3A.
FT CHAIN 1457 1543 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1544 1565 PICORNAIN 3C.
FT CHAIN 1566 1747 RNA-DIRECTED RNA POLYMERASE P3D.
FT CHAIN 1748 2209 MYRISTATE.
FT LIPID 2 2 MYRISTATE.
FT ACT_SITE 1712 1712 PROTEASE (POTENTIAL).
FT ACT_SITE 1726 1726 PROTEASE (POTENTIAL).
SQ SEQUENCE 2209 AA; 246576 MW; 9EC1FF4D174A28A4 CRC64;

Query Match 4.7%; Score 137.5; DB 1; Length 2209;
Best Local Similarity 19.3%; Pred. No. 0.036;
Matches 108; Conservative 73; Mismatches 193; Indels 187; Gaps 25;

QY 22 VPEANTAEPISEMPVAGATAAATAGOVNMIDPWIMNNVYVQAPQGEFTISPNNTP--GDI 79
DB 366 LPEFDVTPPID---IPGEVKNMELAEIDTMIPLNLENTKRTMDMYRVRLSDKPHITDDP 422
QY 80 LFDLQLGPHLNPFLSH-----LAQMYNGWGNMKVKVLLAGNAFTAGKTIISCIIPGFAA 134
```

Search completed: January 16, 2003, 09:55:17
Job time : 23 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 09:52:25 ; Search time 38 Seconds
(without alignments)
2955.149 Million cell updates/sec

Title: US-09-926-799-1

Perfect score: 2896

Sequence: 1 MMASKDAPTNMDGTSGAGQ.....YOLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2873	99.2	545	12 Q8VA02	Q8VA02 human calic
2	2846	98.3	545	12 Q9WI82	Q9WI82 norwalk-lik
3	2846	98.3	545	12 Q918C8	Q918C8 norwalk-lik
4	2840	98.1	545	12 Q918C2	Q918C2 norwalk-lik
5	2798	96.6	545	12 Q91H12	Q91H12 human calic
6	2744	94.8	545	12 Q9IV45	Q9IV45 human calic
7	2605.5	90.0	544	12 Q66418	Q66418 desert shie
8	2124	73.3	539	12 Q9IV48	Q9IV48 human calic
9	1979	68.3	544	12 Q9DU46	Q9DU46 chiba virus
10	1975	68.2	544	12 Q91I85	Q91I85 human calic
11	1974	68.2	544	12 Q9QTE7	Q9QTE7 chiba virus
12	1972.5	68.1	540	12 Q9VS14	Q9VS14 norwalk vir
13	1968	68.0	544	12 Q918C5	Q918C5 norwalk-lik
14	1964.5	67.8	546	12 Q9IV47	Q9IV47 human calic
15	1956.5	67.6	543	12 Q9IV43	Q9IV43 human calic
16	1956.5	67.6	544	12 Q8VA27	Q8VA27 human calic

17	1954	67.5	544	12 Q918D1	Q918D1 norwalk-lik
18	1952	67.4	544	12 Q91E32	Q91E32 human calic
19	1946.5	67.2	543	12 Q918B9	Q918B9 norwalk-lik
20	1946	67.2	530	12 Q83884	Q83884 norwalk vir
21	1942	67.1	544	12 Q9IV36	Q9IV36 human calic
22	1941	67.0	544	12 Q9IV41	Q9IV41 human calic
23	1938	66.9	530	12 Q9Q739	Q9Q739 norwalk-lik
24	1930.5	66.7	531	12 Q8V7J5	Q8V7J5 norwalk-lik
25	1908.5	65.9	517	12 Q9JH41	Q9JH41 norwalk-lik
26	1900	65.6	530	12 Q83876	Q83876 norwalk vir
27	1886	65.1	541	12 Q9IV42	Q9IV42 human calic
28	1274	44.0	522	12 Q8V639	Q8V639 norwalk-lik
29	1268	43.8	522	12 Q8V638	Q8V638 norwalk-lik
30	1260.5	43.5	519	12 Q9YQ22	Q9YQ22 bovine cal
31	1228	42.4	540	12 Q917X1	Q917X1 norwalk-lik
32	1223	42.2	542	12 Q66293	Q66293 calicivirus
33	1222	42.2	542	12 Q913B6	Q913B6 human calic
34	1219	42.1	540	12 Q9IV50	Q9IV50 human calic
35	1215.5	42.0	542	12 Q12305	Q12305 snow mounta
36	1213	41.9	542	12 Q96877	Q96877 snow mounta
37	1211	41.8	538	12 Q91I65	Q91I65 human calic
38	1203	41.5	540	12 Q917W8	Q917W8 norwalk-lik
39	1202.5	41.5	535	12 Q915C6	Q915C6 human calic
40	1202.5	41.5	540	12 Q918B0	Q918B0 norwalk-lik
41	1198.5	41.4	548	12 Q88291	Q88291 small round
42	1196	41.3	537	12 Q913B7	Q913B7 human calic
43	1189	41.1	537	12 Q91H09	Q91H09 human calic
44	1188.5	41.0	548	12 Q917Z5	Q917Z5 norwalk-lik
45	1188.5	41.0	548	12 Q916E5	Q916E5 human calic

ALIGNMENTS

RESULT 1

Q8VA02 PRELIMINARY; PRT: 545 AA.
AC Q8VA02;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Potsdam 196/2000/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=150135;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/POTSdam 196/2000/DE;
RA Kuenkel U., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF439267; AAL32455.1;
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 545 AA; 58601 MW; 7DA7D4D9F3521D89 CRC64;

Query Match 99.2%; Score 2873; DB 12; Length 545;
Best Local Similarity 99.1%; Pred. No. 1e-222;
Matches 540; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISMPEVAGAAATAAGQVNMIDPWIMNNY 60
|||||
Db 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISMPEVAGAAATAAGQVNMIDPWIMNNY 60
|||||
Qy 61 VOAPGGEFTISPNNTPGDITLFDLQLGPHLNPFLSHLAQMYNGWVGNMKVLLAGNAFTA 120
|||||
Db 61 VOAPGGEFTISPNNTPGDITLFDLQLGPHLNPFLSHLAQMYNGWVGNMKVLLAGNAFTA 120
|||||
Qy 121 GKIIISCIPPGFAAQNISIAQATMFPHPVIADYRVLEPIEVPLEDVNRVLFHNDNAPTMR 180
|||||
Db 121 GKIIISCIPPGFAAQNISIAQATMFPHPVIADYRVLEPIEVPLEDVNRVLFHNDNAPTMR 180
|||||

QY	181	LVCMLYTPLRASGSSGTDPEVIAGRVLTCTSPDSEFLVPPNVEQTKPFSPNPLN	240		121	GKIIISIPPGFAAQNISIAQATMEPHVIADVRVLEPIEVPLEDRVNVLFHNDNTPTMR	180
Db	181	LVCMLYTPLRASGSSGTDPEVIAGRVLTCTSPDSEFLVPPNVEQTKPFSPNPLN	240		181	LVCMLYTPLRASGSSGTDPEVIAGRVLTCTSPDSEFLVPPNVEQTKPFSPNPLN	240
QY	241	TLSNSRVPSLIKSMWVRDHQGMVQFONGRVTLDCQLGTTPTTSASQLCKIRGSVFHANG	300		181	LVCMLYTPLRASGSSGTDPEVIAGRVLTCTSPDSEFLVPPNVEQTKPFSPNPLN	240
Db	241	TLSNSRVPSLIKSMWVRDHQGMVQFONGRVTLDCQLGTTPTTSASQLCKIRGSVFHANG	300		241	TLSNSRVPSLIKSMWVRDHQGMVQFONGRVTLDCQLGTTPTTSASQLCKIRGSVFHANG	300
QY	301	GNGYNLTELDSGSPYHAFESPAPIGFPDLGECDDHMEASPTTQFNTGDVIRKQINVKQESAF	360		241	TLSNSRVPSLIKSMWVRDHQGMVQFONGRVTLDCQLGTTPTTSASQLCKIRGSVFHANG	300
Db	301	GNGYNLTELDSGSPYHAFESPAPIGFPDLGECDDHMEASPTTQFNTGDVIRKQINVKQESAF	360		301	GNGYNLTELDSGSPYHAFESPAPIGFPDLGECDDHMEASPTTQFNTGDVIRKQINVKQESAF	360
QY	361	APHLGTIQADGLSDSVNTNMIAKLGWSPVSDGHRGVDVPMWIPRYGSTLTEAAQLAPP	420		301	GNGYNLTELDSGSPYHAFESPAPIGFPDLGECDDHMEASPTTQFNTGDVIRKQINVKQESAF	360
Db	361	APHLGTIQADGLSDSVNTNMIAKLGWSPVSDGHRGVDVPMWIPRYGSTLTEAAQLAPP	420		361	APHLGTIQADGLSDSVNTNMIAKLGWSPVSDGHRGVDVPMWIPRYGSTLTEAAQLAPP	420
QY	421	IYPPGFGGAIVFFMSDFPIAHGTNGLSVPCCTIPOEFVTHFVNEQAPTRGEAALLHYLDPD	480		361	APHLGTIQADGLSDSVNTNMIAKLGWSPVSDGHRGVDVPMWIPRYGSTLTEAAQLAPP	420
Db	421	IYPPGFGGAIVFFMSDFPIAHGTNGLSVPCCTIPOEFVTHFVNEQAPTRGEAALLHYLDPD	480		421	IYPPGFGGAIVFFMSDFPIAHGTNGLSVPCCTIPOEFVTHFVNEQAPTRGEAALLHYLDPD	480
QY	481	THRNLEGEFKLYPEGFMTCVPNSSGGTPQTLPLNGVFVFSWVSRYQLKPYGTAGPACRL	540		421	IYPPGFGGAIVFFMSDFPIAHGTNGLSVPCCTIPOEFVTHFVNEQAPTRGEAALLHYLDPD	480
Db	481	THRNLEGEFKLYPEGFMTCVPNSSGGTPQTLPLNGVFVFSWVSRYQLKPYGTAGPACRL	540		481	THRNLEGEFKLYPEGFMTCVPNSSGGTPQTLPLNGVFVFSWVSRYQLKPYGTAGPACRL	540
QY	541	GIRRS 545			481	THRNLEGEFKLYPEGFMTCVPNSSGGTPQTLPLNGVFVFSWVSRYQLKPYGTAGPACRL	540
Db	541	GIRRS 545			541	GIRRS 545	
RESULT 3							
ID	Q918C8	PRELIMINARY;	PRT;	545 AA.			
AC	Q918C8;						
DT	01-NOV-1999 (TrEMBLrel. 12, Created)						
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)						
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)						
DE	Capsid protein.						
OS	Norwalk-like virus.						
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;						
OC	Norwalk-like viruses.						
OX	NCBI_TaxID=95340;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=HU/NLV/STAV/95/NOR;						
RX	MEDLINE=20349239; PubMed=10893150;						
RA	Myrmel M., Rimstad E.;						
RT	"Antigenic diversity of Norwalk-like viruses: expression of the capsid						
RT	protein of a genogroup I virus, distantly related to Norwalk virus.,"						
RL	Arch. Virol. 145:711-723(2000).						
RN	[2]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=HU/NLV/STAV/95/NOR;						
RA	Myrmel M., Rimstad E.;						
RL	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.						
DR	EMBL; AF145709; AAD37377.1; -						
DR	InterPro; IPR004005; Calici.coat.						
DR	Pfam; PF00915; Calici_coat; 1.						
SQ	SEQUENCE 545 AA; 58631 MW; 89D035D4FEB10BEF CRC64;						
Query Match 98.3%; Score 2846; DB 12; Length 545;							
Best Local Similarity 98.3%; Pred. No. 1.6e-220;							
Matches 536; Conservative 4; Mismatches 5; Indels 0; Gaps 0;							
QY	1	MMASKDAPTNMDTSGAGOLVPEANTAEPTSMPEVAGAAATAAGVNMIDPWIMNNY	60		1	MMASKDAPTNMDTSGAGOLVPEANTAEPTSMPEVAGAAATAAGVNMIDPWIMNNY	60
Db	1	MMASKDAPTNMDTSGAGOLVPEANTAEPTSMPEVAGAAATAAGVNMIDPWIMNNY	60		1	MMASKDAPTNMDTSGAGOLVPEANTAEPTSMPEVAGAAATAAGVNMIDPWIMNNY	60
QY	61	VOAPOGETISPNNPTGGDILDLQLGPHLNPFLSLHAQYNGWGNMKVKVLLAGNAFTA	120		61	VOAPOGETISPNNPTGGDILDLQLGPHLNPFLSLHAQYNGWGNMKVKVLLAGNAFTA	120
Db	61	VOAPOGETISPNNPTGGDILDLQLGPHLNPFLSLHAQYNGWGNMKVKVLLAGNAFTA	120		61	VOAPOGETISPNNPTGGDILDLQLGPHLNPFLSLHAQYNGWGNMKVKVLLAGNAFTA	120
QY	121	GKIIISIPPGFAAQNISIAQATMEPHVIADVRVLEPIEVPLEDRVNVLFHNDNTPTMR	180		121	GKIIISIPPGFAAQNISIAQATMEPHVIADVRVLEPIEVPLEDRVNVLFHNDNTPTMR	180

OS Human calicivirus NLV/VA98115/1998.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=165509;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/VA98115/1998;
RA Jiang X., Zhong W.M., Farkas T., Huang P.W., Wilton N., Barrett E.,
RA Fulton D., Morrow R., Watson D.O.;
RA "Baculovirus expression and antigenic characterization of the capsid
RT proteins of three Norwalk-like viruses.";
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AY038598; AAK84673.1; -;
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat.1;
DR SEQUENCE 545 AA; 58295 MW; 259B0C7DEDB34DAC CRC64;
SQ

Query Match 96.6%; Score 2798; DB 12; Length 545;
Best Local Similarity 96.7%; Pred. NO. 1.le-216;
Matches 527; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 MMWASKDAPTNMDGTSAGOLVPEANTAEPIISMEPVAGATAAATAGOVNMIDPWIMNNY 60
Db 1 MMWASKDAPTNMDGTSAGOLVPEANTAEPIISMEPVAGATAAATAGOVNMIDPWIMNNY 60

QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
Db 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120

QY 121 GKIIISCIPPGFAAQNISIAQATMFPHVIADRVLEPIEVPLEDVNRVLFHNNDAPTMR 180
Db 121 GKIIISCIPPGFAAQNISIAQATMFPHVIADRVLEPIEVPLEDVNRVLFHNNDAPTMR 180

QY 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSEFLFLVPPNVEQTKPFSPVNLPLN 240
Db 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSEFLFLVPPNVEQTKPFSPVNLPLN 240

QY 241 TLSNSRVPSLIRSMVSRDHGMQVQFNGRVTLDGQLGGTTPTSASOLCKIRGSVFHANG 300
Db 241 TLSNSRVPSLIRSMVSRDHGMQVQFNGRVTLDGQLGGTTPTSASOLCKIRGSVFHANG 300

QY 301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGECDDHMEASPTTQFNTGDIKQINVKQESAF 360
Db 301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGECDDHMEASPTTQFNTGDIKQINVKQESAF 360

QY 361 APHLGTTQADGLSDVSNTNMIKLGWVSPVSDGHRGDDVPWIPRYGSTLTEAAQLAPP 420
Db 361 APHLGTTQADGLSDVSNTNMIKLGWVSPVSDGHRGDDVPWIPRYGSTLTEAAQLAPP 420

QY 421 IYPPGFGAIVFFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNEQAPTRGEAALLHYLDPD 480
Db 421 IYPPGFGAIVFFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNEQAPTRGEAALLHYLDPD 480

QY 481 THRNLFGEFKLYPEGFMTCPVNSSGTGPTLPINGVFVSVWSRFYQLKPVGTAGPACRL 540
Db 481 THRNLFGEFKLYPEGFMTCPVNSSGTGPTLPINGVFVSVWSRFYQLKPVGTAGPACRL 540

QY 541 GIRRS 545
Db 541 GIRRS 545

RESULT 6
Q9IV45 PRELIMINARY; PRT; 545 AA.
ID Q9IV45
AC Q9IV45;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Birmingham/93/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.

OX NCBI_TaxID=122916;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/BIRMINGHAM/93/UK;
RX MEDLINE=20404883; PubMed=10949950;
RA Green J., Vinje J., Callimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RA "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL: AJ277612; CAB89093.1; -;
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat.1;
DR SEQUENCE 545 AA; 58689 MW; DC6EBC6EBECDEF7C7 CRC64;
SQ

Query Match 94.8%; Score 2744; DB 12; Length 545;
Best Local Similarity 95.8%; Pred. NO. 2.5e-212;
Matches 522; Conservative 7; Mismatches 16; Indels 0; Gaps 0;

QY 1 MMWASKDAPTNMDGTSAGOLVPEANTAEPIISMEPVAGATAAATAGOVNMIDPWIMNNY 60
Db 1 MMWASKDAPTNMDGTSAGOLVPEANTAEPIISMEPVAGATAAATAGOVNMIDPWIMNNY 60

QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
Db 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120

QY 121 GKIIISCIPPGFAAQNISIAQATMFPHVIADRVLEPIEVPLEDVNRVLFHNNDAPTMR 180
Db 121 GKIIISCIPPGFAAQNISIAQATMFPHVIADRVLEPIEVPLEDVNRVLFHNNDAPTMR 180

QY 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSEFLFLVPPNVEQTKPFSPVNLPLN 240
Db 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSEFLFLVPPNVEQTKPFSPVNLPLN 240

QY 241 TLSNSRVPSLIRSMVSRDHGMQVQFNGRVTLDGQLGGTTPTSASOLCKIRGSVFHANG 300
Db 241 TLSNSRVPSLIRSMVSRDHGMQVQFNGRVTLDGQLGGTTPTSASOLCKIRGSVFHANG 300

QY 301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGECDDHMEASPTTQFNTGDIKQINVKQESAF 360
Db 301 GNGYNLTLDGSPYHAFESPAPIGFDPDLGECDDHMEASPTTQFNTGDIKQINVKQESAF 360

QY 361 APHLGTTQADGLSDVSNTNMIKLGWVSPVSDGHRGDDVPWIPRYGSTLTEAAQLAPP 420
Db 361 APHLGTTQADGLSDVSNTNMIKLGWVSPVSDGHRGDDVPWIPRYGSTLTEAAQLAPP 420

QY 421 IYPPGFGAIVFFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNEQAPTRGEAALLHYLDPD 480
Db 421 IYPPGFGAIVFFMSDFPIAHGTNGLSVPCPTIPOEFVTHFVNEQAPTRGEAALLHYLDPD 480

QY 481 THRNLFGEFKLYPEGFMTCPVNSSGTGPTLPINGVFVSVWSRFYQLKPVGTAGPACRL 540
Db 481 THRNLFGEFKLYPEGFMTCPVNSSGTGPTLPINGVFVSVWSRFYQLKPVGTAGPACRL 540

QY 541 GIRRS 545
Db 541 GIRRS 545

RESULT 7
Q66418 PRELIMINARY; PRT; 544 AA.
ID Q66418
AC Q66418;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Capsid protein.
OS Dessert Shield virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=33755;
RN [1]
RP SEQUENCE FROM N.A.


```
RC STRAIN=DSV2395;
RA MEDLINE=94174735; PubMed=8128635;
RX Lew J.F., Kapikian A.Z., Jiang X., Estes M.K., Green K.Y.;
RT "Molecular characterization and expression of the capsid protein of a
RT Norwalk-like virus recovered from a Desert Shield troop with
RT gastroenteritis.";
RL Virology 200:319-325(1994).
DR EMBL: U04469; AAA16285.1; -.
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58606 MW; 07A4AE7E4BCE81E8 CRC64;

Query Match 90.0%; Score 2605.5; DB 12; Length 544;
Best Local Similarity 89.2%; Pred. No. 3.5e-201;
Matches 486; Conservative 24; Mismatches 34; Indels 1; Gaps 1;

QY 1 MMASKDAPNTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNNY 60
DB 1 MMASKDAPNTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNNY 60
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
QY 121 GKIIISCIPPGFAAQNISIAQATMPHFVIADVRVLEPIEVPLEDVNRVLFHNNNDAPTMR 180
DB 121 GKIIISCIPPGFAAQNISIAQATMPHFVIADVRVLEPIEVPLEDVNRVLFHNNNDAPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVBEQTKPFESVPLNPLN 240
DB 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVBEQTKPFESVPLNPLN 240
QY 241 TLSNSRVPSLIKSMVSRDHGMQVQFQNGRVTLDGLOGQTTPTSASQLCKIRGSVPHANG 300
DB 241 TLSNSRVPSLIKSMVSRDHGMQVQFQNGRVTLDGLOGQTTPTSASQLCKIRGSVPHANG 300
QY 301 GNGYNLTLDGSPYHAFESPAPIGPDPDLGECDDWHM--EASPTQFNTGDVQKINVKQESAF 360
DB 301 GNGYNLTLDGSPYHAFESPAPIGPDPDLGECDDWHM--EASPTQFNTGDVQKINVKQESAF 360
QY 361 APHLGTIOADGLSDSVNTNMIAKLGWSPVSDGHRGDVDPWVIPRYGSTLTEAAQLAPP 420
DB 361 APHLGHVQADNLIS-AGANTDLIVLSWISVSDQHRHDVDPWVIPRYGSTLTEAAQLAPP 419
QY 421 IYPPGGEAIVFMSDFPIAHGTNGLSVPTCTIQEFVTHFVNEQAPTRGEAALLHYLDPD 480
DB 421 IYPPGGEAIVFMSDFPIAHGTNGLSVPTCTIQEFVTHFVNEQAPTRGEAALLHYLDPD 480
QY 481 THRNLFGEFKLYPEGFMTCVPNSGSGTGPQTLPIGVFVSWVSRYQLKPVGTAGPACRL 540
DB 481 THRNLFGEFKLYPEGFMTCVPNSGSGTGPQTLPIGVFVSWVSRYQLKPVGTAGPACRL 539
QY 541 GIRR 545
DB 540 GIRR 544

RESULT 8
Q91V48 PRELIMINARY; PRT: 539 AA.
AC Q91V48:
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Winchester/94/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=122913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/WINCHESTER/94/UK;
RX MEDLINE=20404883; PubMed=10949950;
```

```
RA Green J., Vinje J., Callimore C.I., Koopmans M., Hale A.D.,
RT Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL: AJ277609; CAB89090.1; -.
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58120 MW; 14E612D7A576395D CRC64;

Query Match 73.3%; Score 2124; DB 12; Length 539;
Best Local Similarity 72.7%; Pred. No. 1.9e-162;
Matches 397; Conservative 56; Mismatches 83; Indels 10; Gaps 6;

QY 1 MMASKDAPNTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNNY 60
DB 1 MMASKDAPNTNMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNNY 60
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
QY 121 GKIIISCIPPGFAAQNISIAQATMPHFVIADVRVLEPIEVPLEDVNRVLFHNNNDAPTMR 180
DB 121 GKIIISCIPPGFAAQNISIAQATMPHFVIADVRVLEPIEVPLEDVNRVLFHNNNDAPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVBEQTKPFESVPLNPLN 240
DB 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVBEQTKPFESVPLNPLN 240
QY 241 TLSNSRVPSLIKSMVSRDHGMQVQFQNGRVTLDGLOGQTTPTSASQLCKIRGSVPHANG 300
DB 241 TLSNSRVPSLIKSMVSRDHGMQVQFQNGRVTLDGLOGQTTPTSASQLCKIRGSVPHANG 300
QY 301 GNGYNLTLDGSPYHAFESPAPIGPDPDLGECDDWHM--EASPTQFNTGDVQKINVKQES 358
DB 301 GNGYNLTLDGSPYHAFESPAPIGPDPDLGECDDWHM--EASPTQFNTGDVQKINVKQES 358
QY 359 AFAPHLGTIOADGLSDSVNTNMIAKLGWSPVSDGHRGDVDPWVIPRYGSTLTEAAQLA 418
DB 359 PICTXLGSIET--TSQDQDPAQDGLGTLAWVSPSTSCAR--VDPWKIPSYGSTVTESTILA 414
QY 419 PIYPPGGEAIVFMSDFPIAHGTNGLSVPTCTIQEFVTHFVNEQAPTRGEAALLHYLD 478
DB 419 PIYPPGGEAIVFMSDFPIAHGTNGLSVPTCTIQEFVTHFVNEQAPTRGEAALLHYLD 473
QY 479 PDTHRNLFGEFKLYPEGFMTCVPNSGSGTGPQTLPIGVFVSWVSRYQLKPVGTAGPAC 538
DB 479 PDTHRNLFGEFKLYPEGFMTCVPNSGSGTGPQTLPIGVFVSWVSRYQLKPVGTAGPAC 532
QY 539 GLIRR 544
DB 533 GLGVR 538

RESULT 9
Q9DU46 PRELIMINARY; PRT: 544 AA.
AC Q9DU46:
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=99565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/CHIBA 407/1987/JP;
RA Someya Y., Takeda N., Miyamura T.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=HU/NLV/CHIBA 407/1987/JP;
RX MEDLINE=20569531; PubMed=11118371;
RA Someya Y., Takeda N., Miyamura T.;
RT "Complete nucleotide sequence of the Chiba virus genome and functional
RL expression of the 3C-like protease in Escherichia coli.";
RL Virology 278:490-500(2000).
DR EMBL; AB042808; BAB18267.1; -.
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58351 MW; 1DB7A4E81FBB246F CRC64;

Query Match 68.3%; Score 1979; DB 12; Length 544;
Best Local Similarity 66.8%; Pred. No. 9e-151;
Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;

QY 1 MMASKDAPTNMDGTSAGOLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNNY 60
DB 1 MMASKDAPTSADGATGAGOLVPEVNTADIPIDPVAGSSTALATAGQVNLDPWIINF 60
QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGMKVRVVLAGNAFTA 120
QY 121 GKIIISCIPPFQAQNIATAQATMFPHVIADRVRLPEIEVPLEDVRLVFNHNDNAPTMR 180
DB 121 GKVIICVPPGFSQRTLSIAQAATLFPHVIAVRTLDPEVEPLEDVRLVFNHNDTQPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQKTKPFSVPLNPLN 240
DB 181 LVCMLYTPLRASGSGTDSFVAGRVLTCPGPDFNFLEFLVPPVEQKTRPFTVPNIPLK 240
QY 241 TLSNSRVPSLIKSMVSRDHGMQVQNGRVTLTGLOGTTPTTSASOLCKIRGSVFHANG 300
DB 241 YLSNSRIPNPIEGMSLSPDQTNVQFNGRCTIDGQPLGTPVSVSOLCKFRGRI--TSG 298
QY 301 GNGYNLTDELGSPHAFESAPITGFPDLGECDDHMEAS--PTTQFNTGDVIKQINVKQES 358
DB 299 QRVNLNLTDELGSPMAFAAPAGFPDLGSCDWHIEMSKIPNSSTQNNPIVTNSVKPNSQ 358
QY 359 AFAPHLGTQADGLSDVSNTNMIAKLGWSPVSDGHRGVDVPIPRYGSTLTEAAOLA 418
DB 359 QFVPHLSITLID--ENVSSGGDYIGTQWTSPPSDSGGANTFNWKIPDYGSSLAESAOLA 416
QY 419 PPIYPGGEAIVFMSDFPIAHGTNGLS----VPCIPQEFVTHFVNEQAPTRGEAALL 474
DB 417 PAVYPPGGEVIVFMAIP---GPNQSGSPNLVPCLLPQEVITHFISEQAPIQGEAALL 473
QY 475 HYLDPDTHRNLTGERFKLYPEGMTCPVNSSGTGPQTLPLNGVVFVSWVSRFYQLKPVGTA 534
DB 474 HYVDPDTNRNLGERFKLYPGGYLTCPVNSSSTGPQQLPLDGVFVFASWVSRFYQLKPVGTA 533
QY 535 GPA-CRLGIRR 544
DB 534 GPARGRLGVR 544

RESULT 10
Q91185 ID Q91185 PRELIMINARY; PRT: 544 AA.
AC Q91185;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Koblenz/433/2000/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=165252;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/KOBLENZ 433/2000/DE;
RA Kuenkel U., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
```

```
RT with Norwalk-like viruses in Germany.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF394960; AAK72048.1; -.
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58222 MW; 1B39A595B733A68 CRC64;

Query Match 68.2%; Score 1975; DB 12; Length 544;
Best Local Similarity 66.8%; Pred. No. 1.9e-150;
Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;

QY 1 MMASKDAPTNMDGTSAGOLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNNY 60
DB 1 MMASKDAPTSADGATGAGOLVPEVNTAGPIIDPVAGSSTALATAGQVNLDPWIINF 60
QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGMKVRVVLAGNAFTA 120
QY 121 GKIIISCIPPFQAQNIATAQATMFPHVIADRVRLPEIEVPLEDVRLVFNHNDNAPTMR 180
DB 121 GKVIICVPPGFSQRTLSIAQAATLFPHVIAVRTLDPEVEPLEDVRLVFNHNDTQPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVEQKTKPFSVPLNPLN 240
DB 181 LVCMLYTPLRASGSGTDSFVAGRVLTCPGPDFNFLEFLVPPVEQKTRPFTVPNIPLK 240
QY 241 TLSNSRVPSLIKSMVSRDHGMQVQNGRVTLTGLOGTTPTTSASOLCKIRGSVFHANG 300
DB 241 YLSNSRIPNPIEGMSLSPDQTNVQFNGRCTIDGQPLGTPVSVSOLCKFRGRI--TSG 298
QY 301 GNGYNLTDELGSPHAFESAPITGFPDLGECDDHMEAS--PTTQFNTGDVIKQINVKQES 358
DB 299 QRVNLNLTDELGSPMAFAAPAGFPDLGSCDWHIEMSKIPNSSTQNNPIVTNSVKPNSQ 358
QY 359 AFAPHLGTQADGLSDVSNTNMIAKLGWSPVSDGHRGVDVPIPRYGSTLTEAAOLA 418
DB 359 QFVPHLSITLID--DNVSSGGDYIGTQWTSPPSDSGGANTFNWKIPDYGSSLAESAOLA 416
QY 419 PPIYPGGEAIVFMSDFPIAHGTNGLS----VPCIPQEFVTHFVNEQAPTRGEAALL 474
DB 417 PAVYPPGGEVIVFMAIP---GPNQSGSPNLVPCLLPQEVITHFISEQAPIQGEAALL 473
QY 475 HYLDPDTHRNLTGERFKLYPEGMTCPVNSSGTGPQTLPLNGVVFVSWVSRFYQLKPVGTA 534
DB 474 HYVDPDTNRNLGERFKLYPGGYLTCPVNSSSTGPQQLPLDGVFVFASWVSRFYQLKPVGTA 533
QY 535 GPA-CRLGIRR 544
DB 534 GPARGRLGVR 544

RESULT 11
Q9QTE7 ID Q9QTE7 PRELIMINARY; PRT: 544 AA.
AC Q9QTE7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=99565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHIBA 407;
RA Natori K., Suzuki K., Yamakawa Y., Tatsumi M., Sakae K., Kobayashi S.,
RA Shinozaki K., Ishiko H., Miyamura T., Takeda N.;
RA "Expression and self-assembly of capsid proteins of the Chiba virus, a
RT genetically distinct Norwalk-like virus.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022679; BAA82106.1; -.

```

DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat.1.
SQ SEQUENCE 544 AA: 58323 MW: ABD1CLCF93D872 CRC64;

Query Match 68.2%; Score 1974; DB 12; Length 544;
Best Local Similarity 66.6%; Pred. No. 2.3e-150;
Matches 367; Conservative 68; Mismatches 102; Indels 14; Gaps 6;

Qy 1 MMASKDAPTNWDGSGAGQLVPEANTAEPIISMPEVAGAAATAAGVNMIDPWIMNNY 60
Db 1 MMASKDAPTSADGATGAGQLVPEVNTADPIIDPVAGSALATAGVNLIDPWIMNF 60
Qy 61 VQAPGEEFTISPNNTPGDILFDLQGLPHLNFSLHQAQYNGVGNMVKVLLAGNAFTA 120
Db 61 VQAPGEEFTISPNNTPGDILFDLQGLPHLNFSLHQAQYNGVGNMVKVLLAGNAFTA 120
Qy 121 GKIIISCIPPGFAAQNISIAQATMPHVIADRVLETEVPLEDVRNVLFNHNDNAPTM 180
Db 121 GKVIICVPPGFGSRTLSIAQATLPHVIADRVLDPEVPLEDVRNVLFNHNDQPTMR 180
Qy 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPPNVVEOKTKPFSVPLN 240
Db 181 LVCMLYTPLRASGSSGTDSPFVAGRVLTCPSPDFSLFLVPPNVVEOKTRPPTVNIPLK 240
Qy 241 TLSNSRVPSLIKSMVSRDHQGMVFQNGRVTLQGLQGTTPTSASQLCKIRGSVFHANG 300
Db 241 YLSNSRIENPNEGMSLDPQTNQVQFQNGRGTIDGQPLGTPVSVSOLCKFRGTI--TSG 298
Qy 301 GNGVNLTELDGSPYHAFESPAPIGPDLGECGDWHMEAS--PTQENTGDVTKQINVKQES 358
Db 299 QRVNLTELDGSPFGFAPAPAGEPDLGSCDWHIEMSKIPNSSTQNNPIVNSVKPNSQ 358
Qy 359 AFAPHLGTIQADGLSDVSVNTNMIAKLQWSPVSDGHRGVDPPWIPRYGSLTLEAAQLA 418
Db 359 QFVPHLSITILD--ENVSSGGYIGTIQWTPSPDSGGANTNEWKIPDYGGSLAEASQLA 416
Qy 419 PPIYPGGEAIVFMSDFPIAHGTNGLS----VPCTTPQEVTHFVNEQAPTRGEAALL 474
Db 417 FAVYPPGNEVIVFMAIP---GNQSGSNLVPCLLPQEQYTHIFISEQAPIQGEAALL 473
Qy 475 HYLDPTHNLNLFGEFLKYPGEGMTCPVNSSGTGPTPLINGVFVFSWVSRYQLKPVGTA 534
Db 474 HYVDPDTNRLNLFGEFLKYPGEGYTCVPSNSSTGTPQPLDGVFVFSWVSRYQLKPVGTA 533
Qy 535 GPA-CRLGIRR 544
Db 534 GPARGRLGVR 544

RESULT 12
Q9YS14 PRELIMINARY; PRT; 540 AA.
AC Q9YS14;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Capsid protein.
OS Norwalk virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=11983;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NORWALK-LIKE;
RX MEDLINE=20254531; PubMed=10795514;
RA Schreier E., Doering F., Kuenkel U.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
RT with small round structured viruses in Germany in 1997/98.";
RL Arch. Virol. 145:443-453(2000).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=NORWALK-LIKE;
RA Schreier E., Doering F., Kuenkel U.;

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF093797; AAC64603.1;
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat.1.
SQ SEQUENCE 540 AA: 58018 MW: 647843487A654CBE CRC64;

Query Match 68.1%; Score 1972.5; DB 12; Length 540;
Best Local Similarity 67.7%; Pred. No. 3e-150;
Matches 377; Conservative 60; Mismatches 89; Indels 31; Gaps 7;

Qy 1 MMASKDAPTNWDGSGAGQLVPEANTAEPIISMPEVAGAAATAAGVNMIDPWIMNNY 60
Db 1 MMASKDAPPSDAGSGAGQLVPEVNTADQISMDPVAGATAVATAGVNMIDPWIMNF 60
Qy 61 VQAPGEEFTISPNNTPGDILFDLQGLPHLNFSLHQAQYNGVGNMVKVLLAGNAFTA 120
Db 61 VQAPGEEFTISPNNTPGDILFDLQGLPHLNFSLHQAQYNGVGNMVKVLLAGNAFTA 120
Qy 121 GKIIISCIPPGFAAQNISIAQATMPHVIADRVLETEVPLEDVRNVLFNHNDN-APTM 179
Db 121 GKVIICVPPGFGDARILTAQATLPHVIADRVLETEVPLEDVRNVLHNSSQPTM 180
Qy 180 RLVCMLYTPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFLVPPNVVEOKTKPFSVPLN 239
Db 181 RLVCMLYTPLRASGSSGTDAPFVAGRVLTCPSPDFSLFLVPPNVVEOKTRVFSVPNIPL 240
Qy 240 NTLNSRVPSLIKSMVSRDHQGMVFQNGRVTLQGLQGTTPTSASQLCKIRGSVFHANG 299
Db 241 KDLNSRVPTLQGMFVSPDVNQSVQFQNGRCQIDGQLOGTTPVLSQLCKIRGT--SS 298
Qy 300 GNGVNLTELDGSPYHAFESPAPIGPDLGECGDWHM-----EASPTTQNTGQDVTK 350
Db 299 NTRVLNLSVDGTPPEVPLESPAPVGPDPDGGDWHVGETFEARDQDPSONVTEATND--- 355
Qy 351 QINVQESAFAPHLGTI--QADGLSDVSVNTNMIAKLQWSPVSDGHRGVDPPWIPRY 407
Db 356 -----SSEVYLGSI SPHNGDGFH----SCDIIGSLDOWISAPSDGSALDV--WSIPKY 402
Qy 408 GSTLEAQAAPPYPPGFEAIVFMSDFPIAHGTNGLSVPCCTIPQEFVTHFVNEQAPT 467
Db 403 GSSLPDVTHLAPAVPPGFGFEVILYFHSKFGSGPTDKLRVPCLLPQEQYTHIFCNEAPI 462
Qy 468 RGEAALLHYLDPTHNLNLFGEFLKYPGEGMTCPVNSSGTGPTPLINGVFVFSWVSRYQL 527
Db 463 AGEAALLHYVDPTDGRNLGFEFLKYPGEGMTCPVNSISSGPTPLINGVFVFSWVSRYQL 522
Qy 528 LKPVGTAGPACRLGIRR 544
Db 523 LKPVGTASARRLGIRR 539

RESULT 13
Q918C5 PRELIMINARY; PRT; 544 AA.
AC Q918C5;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Baltimore/277/1993/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=171836;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/BALTIMORE/277/1993/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).

```

[2]
RN SEQUENCE OF 100-192 FROM N.A.
RC STRAIN-HU/NLV/BALTIMORE/277/1993/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383(1997).
[3]
RN SEQUENCE FROM N.A.
RC STRAIN-HU/NLV/BALTIMORE/277/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.';
RL J. Infect. Dis. 181:S336-S348(2000).
[4]
RN SEQUENCE FROM N.A.
RC STRAIN-HU/NLV/BALTIMORE/277/1993/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414404; AAL12965.1; -.
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA: 58208 MW: D5170FA4B958A672 CRC64;

Query Match 68.0%; Score 1968; DB 12; Length 544;
Best Local Similarity 66.4%; Pred. No. 6.9e-150;
Matches 366; Conservative 69; Mismatches 102; Indels 14; Gaps 6;

QY 1 MMASKADPTNMDGTSCAGOLVPEANTAEPI SMEPVAGAATAAATAGVNMIDPWIMNMY 60
DB 1 MMASKADPTSDAGATGAGOLVPEYNTADPI DVPAGSSTALATAGRVNLDPIWINEF 60

QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHNLPFLSHLAQMYNGWGNMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNTPGDILFDLQGLPHNLPFLSHLSQMYNGWGNRVVVLAGNAFTA 120

QY 121 GKIITCIPGFAAONISIAQATMFPHVIADRVRLPIEVPLEDVRNVLFNHNDNAPTMR 180
DB 121 GKILICVCPGFSRVLSTAAQATLFFHVIADRVRLDVPVEVPLEDVRNVLHNNDTQPTMR 180

QY 181 LVCMLYPLRASGSSGCTDFVIAGRVLTCPSPDFSEFLVLPNVEQKTKPFSVNPPLN 240
DB 181 LVCMLYPLRTGGASGCTDFVIVSGRVLTCPGPDFNLFVLPPTVEQKTRPFTVPNPLK 240

QY 241 TLNSRVPSLTKSMVSRDHQMVQFQNGRVTLDGQLGTTPTTSASQLCKIRGSVFHANG 300
DB 241 TLNSRVPSLTKSMVSRDHQMVQFQNGRVTLDGQLGTTPTTSASQLCKIRGSVFHANG 300

QY 301 GNGYNLTLDGSPHAFESPAPIGFDLGCEDWHMEASPT--PTQFNTGDVVKQINVKQES 358
DB 301 GNGYNLTLDGSPHAFESPAPIGFDLGCEDWHMEASPT--PTQFNTGDVVKQINVKQES 358

QY 359 QGVNLTLTLDGSPHAFAPAGFDLGCEDWHIEMSKIIPNSSTQNNPTIVVNSVKNPSQ 358
DB 359 QGVNLTLTLDGSPHAFAPAGFDLGCEDWHIEMSKIIPNSSTQNNPTIVVNSVKNPSQ 358

QY 359 AFAPHLGTIQADGLSDSVNTNMTAKLWVSPVSDGHRGDVDPWVIPRYGSTLTEAAQLA 418
DB 359 QFVPHLSLITLD--ENVSSGSDYIGTQWTSPPSPDSGGANTNFWKIPDYGSSLAESAQLA 416

QY 419 PPIYPGGEALVFFMSDFPIAHGTNGLS----VPCITPQEFVTHFVNEQAPTGEAALL 474
DB 417 PAVYPPGGENEVVFMASIP--GPNQSGSPNLVPCLLPQEIYTHFISEQAPIQGEAALL 473

QY 475 HYLDPDTHNLGFEKLYPEGFMTCVPNSGTCGPTQLPLNGVVFVSVSWRYOLKPVGTA 534
DB 474 HYVDDPTNRLNGFEKLYPGGYLTCVPNSSTGTPQQLPLDGVVFVFSWRYOLKPVGTA 533

QY 535 GPA-CRLGIRR 544
DB 534 GPARGRLGVR 544
```

```

RESULT 14
Q9IV47 PRELIMINARY: PRT: 546 AA.
AC Q9IV47;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Whiterose/96/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=122914;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HU/NLV/WHITEROSE/96/UK;
RX MEDLINE=20404883; PubMed=1094950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL; AJ277610; CAB89091.1; -.
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 546 AA: 58737 MW: 501F9D2BE2A77B51 CRC64;

Query Match 67.8%; Score 1964.5; DB 12; Length 546;
Best Local Similarity 66.2%; Pred. No. 1.3e-149;
Matches 365; Conservative 78; Mismatches 95; Indels 13; Gaps 7;

QY 1 MMASKADPTNMDGTSCAGOLVPEANTAEPI SMEPVAGAATAAATAGVNMIDPWIMNMY 60
DB 1 MMASKAGQSDGSSGAGOLVPEVNTADPLPMEPVAGPTTAVATAGVNMIDPWIVNMF 60

QY 61 VQAPQGEFTISPNTPGDILFDLQGLPHNLPFLSHLAQMYNGWGNMKVKVLLAGNAFTA 120
DB 61 VQSPQGEFTISPNTPGDILFDLQGLPHNLPFLSHLSQMYNGWGNRVVRLLAGNAFSA 120

QY 121 GKIITCIPGFAAONISIAQATMFPHVIADRVRLPIEVPLEDVRNVLFNHNDNAPTMR 180
DB 121 GKIIIVCCVPGFTSSSLTIAQATLFFHVIADRVRLPIEMLPVEDVRNVLHYTNDSPQTR 180

QY 181 LVCMLYPLRASGSSGCTDFVIAGRVLTCPSPDFSEFLVLPNVEQKTKPFSVNPPLN 240
DB 181 LVCMLYPLRTGGSGNSDSFVAGRVLTAPSDSEFLVLPPTIEQKTRAEVNPILQ 240

QY 241 TLNSRVPSLTKSMVSRDHQMVQFQNGRVTLDGQLGTTPTTSASQLCKIRGSVFHANG 300
DB 241 TLNSRVPSLTKSMVSRDHQMVQFQNGRVTLDGQLGTTPTTSASQLCKIRGSVFHANG 300

QY 301 GNGYNLTLDGSPHAFESPAPIGFDLGCEDWHMEASPT--TOFNTGDVVKQINVKQES 358
DB 301 GNGYNLTLDGSPHAFESPAPIGFDLGCEDWHMEASPT--TOFNTGDVVKQINVKQES 358

QY 359 ARTLNLTEVDGKPFMAFDSPAPVGFDFGKCDWHNRVSKTPNNTSSGDPMSRSVSVQINVQ 358
DB 359 AFAPHLGTIQADGLSDSVNTNMTAKLWVSPVSDGHRGDVDPWVIPRYGSTLTEAAQLA 418

QY 359 GVPVPHLSLQDEVENHPGTG-DYICTIEMISQSTPLCTDNLWEIPDYGSSLSQAANLA 417
DB 419 PPIYPGGEALVFFMSDFPIAHGTNGLS----VPCITPQEFVTHFVNEQAPTGEAALL 474

QY 418 PVVFPFGGEALVFFVSAPF---GPNRSAPNDVPCLLPQEIYTHFISEQAPTMGDAALL 474
DB 475 HYLDPDTHNLGFEKLYPEGFMTCVPNSGTCGPTQLPLNGVVFVSVSWRYOLKPVGTA 534

QY 475 HYVDDPTNRLNGFEKLYPGGYLTCVPNGVGAGPQQLPLNGVLFVSVWSRYOLKPVGTA 534
DB 535 GPA-CRLGIRR 544
DB 535 STARGRLGVR 545

RESULT 15
Q9IV43 PRELIMINARY: PRT: 543 AA.
ID Q9IV43
```

```
AC O9IV43:
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Musgrove/89/UK.
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norwalk-like viruses.
OX NCBI_TaxID=122918;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/MUSGROVE/89/UK;
RX MEDLINE=20404883; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL: AJ277614; CAB89095.1; -.
DR InterPro: IPR004005; Calici_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 543 AA; 58216 MW; 11F97A3989CC540 CRC64;

Query Match 67.6%; Score 1956.5; DB 12; Length 543;
Best Local Similarity 67.8%; Pred. No. 5.8e-149;
Matches 377; Conservative 65; Mismatches 89; Indels 25; Gaps 10;

QY 1 MMASKDAPTINMDGTSGAGOLVPEANTAEPISMPEVAGATAATAAGQVNMIDPWIMNNY 60
DB 1 MMASKDATPSADGAGAGOLVPEVNNAXPLPLDPVAGASTALATAGQVNMIDPWIFNNF 60
QY 61 VQAPQGEFTISPNTTGDILFDLQGLPHLNPFLSLHAQMYNGWGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNTTGDILFDLQGLPHLNPFLSLHAQMYNGWGMKVKVLLAGNAFTA 120
QY 121 GTIISCIPPGFAAQNIAQAQTMFPHVADRVLEPIEVPLEDVRLVFNHNDNAPTMR 180
DB 121 GKVIICVPPGFSRVLSTAAQTLFFPHIADVRTLEPIELPLEDVRLVFNHNDNAPTMR 180
QY 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFVPPNVEQKTKPFSVNPPLN 240
DB 181 LLCMLYTPLRGSGGTDADFVAGRVLTCPSSDFNLFVLPPTVEQKTRPFSVNPPLQ 240
QY *241 TLSNRVPSLIKSMYSRDHQWQVQNGRVTLDGLOGTTPTSASOLCKIRGSVFHANG 300
DB 241 LLSNRVNPNIQSMVLSPOAQNVQFQNGRCTTDGOLLGTPVSVQILKFRGKV--SAG 298
QY 301 GNGYNLTLDGSPYHAFESPAPITGFPDLGECDDHMEAS-PTTQFNTGDVIKQINVKQESA 359
DB 299 SKVINLTLDGSPFLAFEPAPPTGFPDLGTSWDHIEMSLNSNSQSSGNPILLRDIQPNSS 358
QY 360 -FAPHLGTIQADGLSDVSVNTNMIKLGW-----VSPVSDGHRGDVDPWVPIRYGSTLT 412
DB 359 DFVPHLGVSVTTAIDTAGDT--LGTIQWTSQPSNVTVP-----DYNFTIPQYGSSLA 411
QY 413 EAAQLAPPYPPGFGGAIVFFMSDFP---IAHGTNGLSVPTCIPOEFVTHFVNEQAPTRG 469
DB 412 EASQLAPVVPYPPGFGGAIVFFMSPIGPNHTAKPN--LVPCLLPQEFVTHFVSEQAPSMG 469
QY 470 EALLHYLDPDTHRNIGEFKLYPEGFMTCVPSNSSGTGQOTLPINGVVFVSWVSREYQLK 529
DB 470 EAAHYVDPDTHRNIGEFKLYPEGFITCPVN--GTGPOOLPLNGVVFASWVSREYQLK 527
QY 530 PVGTAGPA-CRLGIRR 544
DB 528 PVGTASSARGRLGVR 543
```

Search completed: January 16, 2003, 09:55:56
Job time : 41 secs

THIS PAGE BLANK (USPTO)